Current Anthropology Volume 43, Supplement, August–October 2002
 0011-3204/2002/43supp-0003$3.00

Time in Action

by Charles Goodwin

One effect of the way in which human action is constituted and shaped within a rich multimodal ecology of sign systems is that participants orient to multiple orders of temporality simultaneously. Within talk-in-interaction, linguistic structure provides resources that can be used simultaneously to (1) structure time in the world being represented through talk and (2) provide hearers with resources for projecting future events in the current and future interactions. Such structure in the stream of speech is framed by the participants' bodies. Through interactively organized gesture and posture, participants display crucial information about the temporal and sequential organization of their joint participation in the current interaction. This multiplicity of concurrently relevant embodied temporalities extends to the tools and documents used in a scientific work setting such as an archaeological excavation. To uncover a past world archaeologists use tools from a professional past (e.g., the coding sheet of a senior investigator, the history of research encapsulated in the Munsell color chart, etc.) to build a work-relevant future (the records that will form the basis for subsequent analysis). The data for the present analysis consist of videotapes of situated human interaction.

CHARLES GOODWIN is Professor of Applied Linguistics at the University of California, Los Angeles (Los Angeles, Calif. 90095-1531 [cgoodwin@humnet.ucla.edu]). He received his Ph.D. from the University of Pennsylvania in 1977 and taught at the University of South Carolina until joining the faculty at UCLA in 1996. His publications include *Conversational Interaction* (New York: Academic Press, 1981), "Transparent Vision," in *Interaction and Grammar*, edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra Thompson (Cambridge: Cambridge University Press, 1996), "Action and Embodiment within Situated Human Interaction" (*Journal of Pragmatics* 32:1489–1522), and the edited volume *Conversation and Brain Damage* (Oxford: Oxford University Press, in press).

Using video and audio tapes recorded in a variety of settings, including archaeological field excavations and storytelling in a range of different groups, this paper will investigate how human beings build, in concert with each other, the actions and activities that define their social worlds. The events that make up these settings are constituted through temporally unfolding sequences of action. Individual actions are constructed through the co-articulation of different kinds of phenomena in different media (for example, talk, the body, documents, tools, etc.), each of them capable of providing quite distinct forms of temporal and sequential organization. The effect of this is that a variety of different frameworks for the organization of time are implicated in even apparently simple actions such as a single utterance.

This perspective on the organization of time in action has methodological as well as theoretical consequences. If the methods used to make a record of an event exclude relevant features of the setting and context (for example, by extracting an utterance from the larger sequence of action within which it emerged or recording the talk in a situated face-to-face interaction while ignoring the bodies of the participants, the materials they are working with, etc.), then various phenomena are rendered invisible for subsequent analysis. This may, of course, be quite appropriate. Not only is it impossible to make a record that includes everything, but a theoretical formulation of what is to be investigated will make some phenomena crucial and others marginal or irrelevant. Record making, with its inevitable selectivity, can clarify the task of analysis by eliminating what is considered distracting noise. However, it has also long been recognized that the way in which a record both selects and shapes what is being recorded can deeply influence analysis in ways that the analyst may not recognize, especially when not all the potentially relevant phenomena are visible. Thus Ochs (1979) argues that transcription is in fact a form of theory, and it has frequently been observed that linguistics as a field of study builds upon a long and complex history of writing practices. These practices have clarified certain phenomena, crucially relevant structure in the stream of speech, while obscuring or rendering completely invisible many contextual features that are central to the production of talk (e.g., the participation of hearers, the organization of the setting). It is therefore necessary to take a reflexive stance with regard to the interplay between methods for recording and transcribing an event, the phenomena that alternative choices reveal or hide, and the kinds of analysis that can then be developed.

To explore these issues I will report briefly on some of my research, published elsewhere (Goodwin 2000*c*), that investigates the practices used by archaeologists to classify the color of the dirt they are excavating. Archaeologists use tools created in the past to construct records for future analysis. Their tools and documents are saturated with methods not only for describing the past being excavated but also for organizing their own work practices within consequential retrospective and prospective frameworks. The multiplicity of frameworks for the organization of time instantiated in the specifics

of diverse media is not, however, unique to specialized domains such as scientific practice. The remainder of the paper will therefore examine a mundane but pervasive form of human social organization: people talking to each other. Structure in the stream of speech provides participants with resources for reporting a past while projecting future action in the current interaction. Records of talk and their analysis focus primarily on this structure. However, in face-to-face interaction talk is framed by the bodies of the participants. Records and analysis that include embodiment as well as talk and phenomena such as the records and tools of the archaeologists reveal that human beings are in fact working with a range of temporal and sequential frameworks as they organize action in a rich multimodal environment.

Classifying Color

The study of the way human beings classify color provides an opportunity to investigate how both theoretical frameworks and the records available to the analyst shape the way a phenomenon can be studied, including whether time is a relevant feature of the analysis. The work by Berlin and Kay (1967, 1969) on color categories is among the most distinguished accomplishments in cognitive anthropology and is an excellent example of one major approach to the study of human cognition. Different languages classify the color spectrum in different ways. This has been argued to provide evidence for the Sapir-Whorf hypothesis that language structures perception of the world (Bruner, Oliver, and Greenfield 1966, Greenfield and Bruner 1966). However, Berlin and Kay (1969) demonstrated that the diversity of human color systems was built on a universal infrastructure, one almost certainly linked to structures in the brain. To show this they first located a basic set of color terms in a number of different languages and then had speakers of those languages show which color patches on a Munsell color chart fell within the boundaries of each basic color term. The Munsell chart, consisting of carefully prepared samples of precisely defined colors arranged in a grid, is the accepted reference standard for color description. When Berlin and Kay compared the Munsell maps for different languages they found that all languages locate the foci of their basic color labels at roughly the same place in the color spectrum and, moreover, that a universal pattern exists for adding basic color terms to a language. If a language has only two color names they will be black and white, if it has three the third will be red, the fourth will be either green or yellow, blue will be added next, etc.

In this research tradition, the structures that provide mechanisms for a central human cognitive task such as color classification reside in two interrelated places: the human brain and a linguistic system. Cognition is a psychological process, and its crucial machinery is found within the human skull. Berlin and Kay never looked at how people use color categories to pursue a relevant course of action in the scenes that make up their lifeworld. Instead, all of their informants were performing exactly the same experimental task, and, many (the Tzeltal-speakers being a notable exception) resided in the San Francisco Bay area. The notion of a community of practitioners engaged in color classification was completely irrelevant to Berlin and Kay's analysis.

However, it is possible to take a different approach. A researcher could locate a group or profession faced with the task of classifying color in order to carry out a consequential course of action and then look ethnographically at how they do it. What practices are used, and what has to be taken into account in order to describe and analyze those practices?

As part of the process of excavating a site, archaeologists are faced with the task of systematically and reliably describing the color of the dirt they are excavating. Many phenomena of interest to archaeologists (for example, the place where a post supporting an ancient structure was located) are primarily visible as color patches in the dirt being excavated. However, 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, including maps, photographs, and coding forms of various types, have to be kept of each stage in the excavation (for more extended analysis see Goodwin 2000*c*, 2000*a*).

For archaeologists the classification of color is accomplished as part of an activity that includes not only color categories but also tools, documents, and systematic work practices. The work of color classification is systematically linked to both other settings and the larger course of work required to excavate the site. The two archaeologists in figure 1 are intently scrutinizing a bit of dirt to determine its color because they have been given a coding form to fill out. The form becomes part of a coherent course of events occurring at quite different times in diverse settings. It was written long before this day's work and, indeed, this field season by a senior archaeologist as part of the process of preparing for excavation. It is designed not only to organize the work of those charged with excavating the site but also to build the documentary record that will be used in the lab to provide an analytic picture of the site, make generalizations, write articles, etc. The task of color classification visible here is thus situated within both retrospective and prospective temporal horizons as the current actors use resources constructed in the past to build records that will form the basis for action in the future.

The form contains spaces for describing the color, consistency, and texture of the dirt being examined. By using it a worker views the world from the perspective it establishes. The form thus shapes in fine detail the work practices and perceptual orientation of the workers excavating the site. Those filling in the form are faced with the task of systematically examining the dirt and making appropriate entries in each slot. Of all the possible ways that the earth could be looked at, their perceptual work is focused on determining the exact color of a minute sample of dirt. They engage in active cognitive work, but

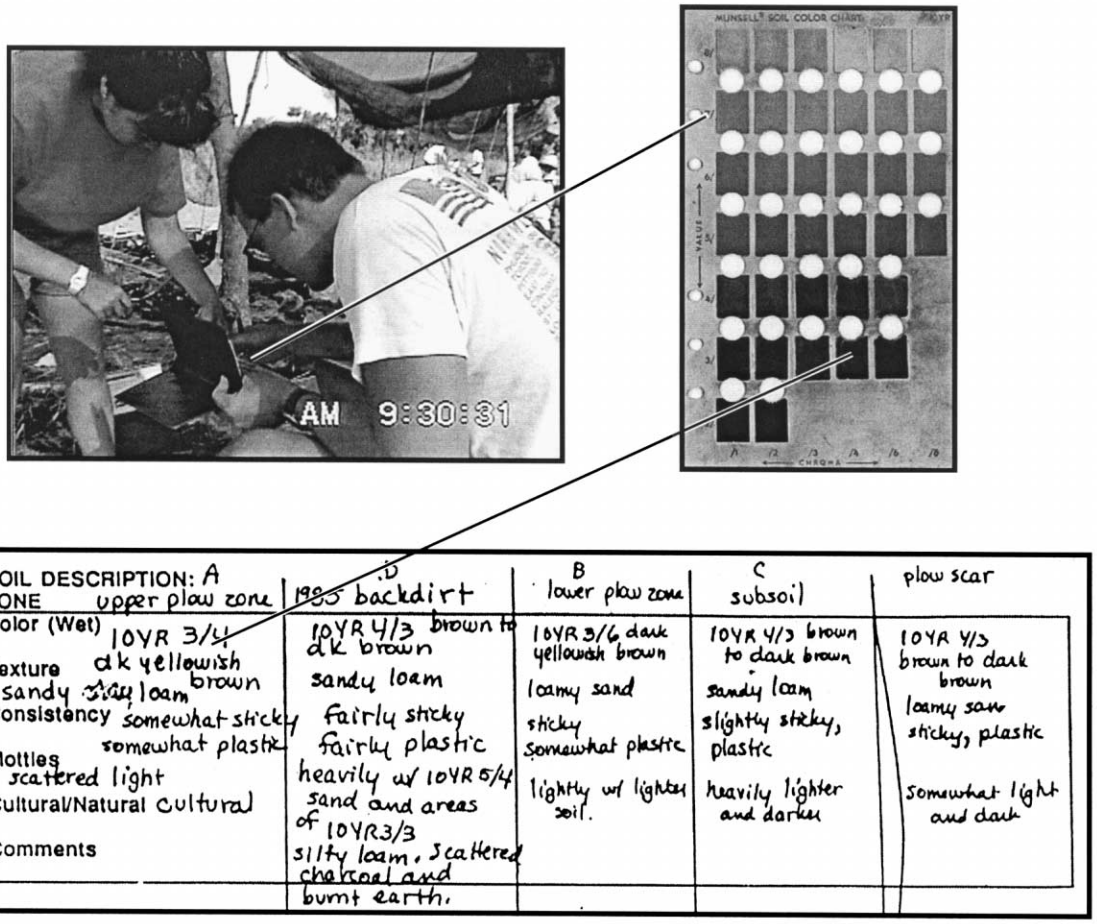


FIG. 1. *Classifying color.*

the parameters of that work have been established by the system that is organizing their perception. Coding schemes distributed on forms allow a senior investigator to inscribe his or her perceptual distinctions into the work practices of the fieldworkers who code the data.

The way in which the coding form brings together on a single page linked actions performed by structurally different kinds of actors at different times and places is shown quite vividly in the contrast between printed text that remains invariant across many settings (for example, category labels such as "color" and "texture" in the document reproduced in figure 1) and the handwritten entries of the different parties who code the data. Under the influence of Bakhtin (1981), linguistic anthropologists have devoted considerable attention to multivocality. However, most of that work has focused on literary genres or oral narrative. Moreover, with the notable exception of work in conversation analysis, dialogic phenomena have been most frequently investigated within the scope of utterances by a single speaker, albeit one reporting the speech of another. A quite different kind of multivocality, one organized by the craft requirements of a work task rather than the genres of the literary academy, can be found in mundane bureaucratic forms. Such documents bring together on a single surface texts produced by different individuals situated at different positions in a work organization.

In order to make an entry in the slot provided for *color*, an archaeologist must make use of another tool, the set of standard color samples provided by a Munsell chart. This chart, with its ordered grid of color samples, incorporates into a portable physical object the results of a long history of scientific investigation into the properties of color. Unlike most documents, the Munsell chart is full of holes (fig. 2). It is not an abstract, generic "text" but a physical document and, moreover, one that makes use of the distinctive properties of paper as a medium. The circular holes cut into the page next to the color patches allow the archaeologist to compare a bit of dirt on a trowel with particular color samples. The trowel is moved from hole to hole until the best fit between the color of the dirt on the trowel and an adjacent patch on the chart is found. With elegant simplicity, the Munsell page, with its holes for viewing the sample of

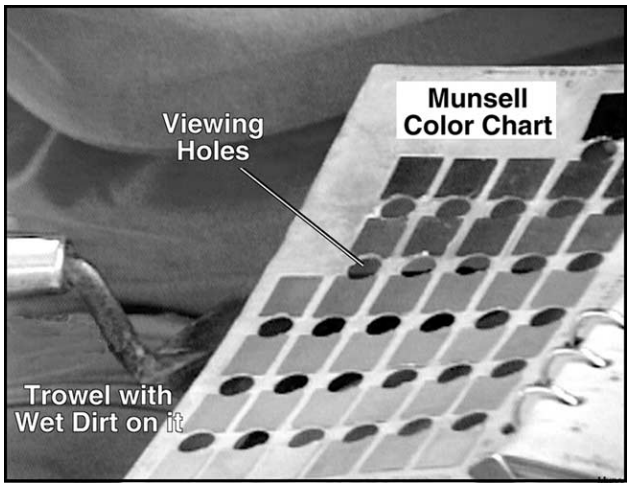


FIG. 2. *Using the Munsell color chart.*

dirt on the trowel, juxtaposes in a single visual field two quite different kinds of spaces: (1) 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 the properties of color, and a physical space instantiated in the orderly modification of variables arranged in a grid on the Munsell page. The pages juxtaposing color patches and viewing holes that allow the dirt to be seen right next to the color sample provide a historically constituted architecture for perception.

As does the excavation form, the Munsell chart both incorporates a history of past action relevant to the present and provides a framework for the accomplishment of action that looks toward the future. Current workers use a tool that encapsulates in a material object theory and solutions developed by earlier workers at other sites faced with the task of color classification. However, rather than simply providing a representation of what has been learned about color, the chart, with its viewing holes next to each color patch, provides a framework for the organization of future action.

Routine work with the Munsell chart seems quite distant from the abstract world of archaeological theory and the debates that are currently animating the discipline. However, the encounter between the coding scheme and the world that occurs as the archaeologist in the field holds a sample of dirt under the Munsell page is one example of a key locus for scientific practice. This is the place where the multifaceted complexity of "nature" is transformed into the phenomenal categories that make up the work environment of a scientific discipline. It is precisely here that nature is transformed into culture.

The work of the archaeologists provides a perspective on color classification that complements that of Berlin and Kay (it does not contradict their work or, indeed, challenge their findings in any way). Rather than taking language categories in isolation as the primary focus of analysis, the work described here focuses on situated practices that encompass semiotic categories, tools, documents, and work-relevant activity systems. This expanded geography of cognition and action has a number of consequences. First, time emerges as a pervasive, constitutive feature of the practices involved in color classification in at least two ways: (1) as a past that shapes in fine detail the environment of the present and the tools used to accomplish work there and (2) as a prospective future toward which the actions in the present are oriented. Both this past and this future are situated within the historically structured work practices and life-world of a group faced with the task of classifying color. Second, rather than drawing an analytic and theoretical division between mental phenomena, such as language categories, and objects in the material world, investigating color as a situated activity system permits the integrated analysis of (1) color categories, (2) tools that provide semiotic structure through the details of their material organization (the ordered grid of color samples provided by the Munsell chart could not exist within a field of action constituted through language alone), and (3) situated practices that shape both the environment being worked in and the way in which that environment is known, classified, and understood.

It might be argued that the work of the archaeologists with the Munsell chart constitutes an esoteric, specialized world. To try to demonstrate that multiple temporalities embodied in diverse media are in fact a general, systematic feature of human action, the remainder of this paper will focus on one of the most pervasive sites for the constitution of human sociality: the organization of talk in face-to-face interaction. It will be argued that in constructing action with talk participants are attending to multiple temporal and sequential frameworks simultaneously. Moreover, these frameworks are embedded in a variety of different media and activity systems including linguistic structure in the stream of speech, prosody, gesture, the sequential organization of action, and the visible body. This inherent diversity makes it possible for new meaning-laden fields, such as the Munsell chart, to be incorporated into what is already, within mundane talk and interaction itself, a multimodal matrix for the organization of human cognition and action.

Projection and Collaborative Action

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 in which their current actions are embedded and relevant phenomena in the world around them. Quite different kinds of temporal and sequential organization interpenetrate this process. Consider, for example, the interactive construction of a state of talk. As a symbolic system, language, through phenomena such as tense and aspect, provides extensive

resources (which vary markedly from language to language) not only for reporting absent events but also for construing them in a variety of temporal relationships to each other. Thus in figure 3, through use of the English past tense ("made," "was") Nancy tells Tasha that both the making of the pie being evaluated and her experience of it occurred at some point prior to the present conversation.[1] In line 3 Tasha also displays enthusiasm for asparagus pie. However, she uses a quite different tense, the habitual present ("I love it"), to formulate her position. Why? Tasha did not actually taste the specific pie that Nancy is talking about. Through use of the habitual present Tasha is able to display appreciation not for the specific pie that Nancy ate but for asparagus pie in general. Thus, rather than merely formulating the time at which events being reported occurred or are expected to occur, tense and aspect provide participants with resources positioning themselves as they collaboratively construct a multiparty evaluation of something in their talk and, moreover, one that displays the differential access that alternative participants have to the entities and events that are the focus of their talk. By making alternative choices from the temporal resources provided by their language, they are able to build in concert with each other an interactively organized stance display—a positive assessment of something—that exhibits each party's congruent appreciation while simultaneously marking their differing access to and resulting perspective with regard to that entity.

1 Nancy: *Jeff* made en asparagus pie.
2 it was *s : : so* ┌*: goo :* d.
3 Tasha: └I love it. °Yeah I love that.

FIG. 3. *Formulating time in talk.*

Projecting Future Action

Grammatical structure for formulating the temporal organization of events being described through language thus constitutes one domain where time structure is relevant to the production of talk. Another, quite different one is provided by the organization of the act of speaking itself. To explore this I will look first at the way in which the timing of Nancy's and Tasha's assessments, specifically their simultaneous production, constructs a particular type of agreement. I will then look at some resources available to these participants for systematically accomplishing such precision timing in conversation (Jefferson 1973). I will argue that in addition to providing resources for formulating the temporal organization of reported events (e.g., that the pie was made and eaten at some point prior to the present), the emerging structure of a strip of talk includes an orientation toward the future. This emerges from the fact that the parties to a state of talk are involved in the collaborative production of social action. In brief, the very same talk that locates the events it is reporting within a specific temporal perspective, such as the past, simultaneously provides a quite different prospective orientation that is implicated in the organization of the utterance as a form of action.

Structurally the display of agreement produced in the data being examined is exceptionally strong. Tasha doesn't wait until after Nancy has produced her evaluation of the pie before providing her own assessment. Instead she overlaps Nancy's "good" with her own "I love it" and thus displays her own position before she has actually heard Nancy's. By acting when she does Tasha runs the risk of positively evaluating something that her coparticipant is taking a negative stance toward (e.g., if Nancy had said "bad" instead of "good" Tasha would have found herself in the position of raving about something that her interlocutor was disparaging). By way of contrast it would be simple to construct agreement by waiting until the second party already knew the position of the first. Through the temporal positioning of their evaluations (e.g., stating their respective positions at exactly the same moment, without having heard the position of the other) these participants are able to construct an elegant demonstration of just how precisely their minds and ways of viewing the world are in tune with each other.

This concurrent assessment, constructed in part through the simultaneous production by separate participants of alternative temporal formulations encoding different experiences of a common entity ("asparagus pie"), each formulation precisely suited to the specific access that each participant has to that entity, provides an example of the accomplishment of elementary human social organization through language use (e.g., shared action collaboratively built by the differentiated work of separate individuals). What practices provide for the systematic possibility of such simultaneous action? More generally, how can a subsequent speaker systematically place her talk precisely at the moment when it is relevant for the accomplishment of joint collaborative action, in this case just at the point when the initial speaker is producing her evaluation of the pie?

A number of different phenomena display emerging temporal structure within a strip of talk. First, the hearable syntax of a speaker's emerging utterance provides a continuously changing set of projections about the kind of unit that can be expected to occur next. In these data the intensifier "so," especially when seen in light of the way in which the prior organization of Nancy's utterance displays that something is being said about Jeff's asparagus pie, creates a syntactic frame within which an assessment adjective is publicly visible as a possible, indeed expected, next unit (e.g., "It was so good/terrible/delicious/ridiculous," etc.). If the hearer wants to

1. Talk is transcribed using the system developed by Gail Jefferson (Sacks, Schegloff, and Jefferson 1974:731–33). A bracket linking lines of talk by separate speakers indicates the point where the talk of the second speaker begins to overlap the talk of the first. Bold italics mark talk that is spoken with special emphasis. Colons indicate that the current sound has been noticeably lengthened.

produce a concurrent assessment of her own she can thus systematically see from the emerging syntax of the talk in progress that this is the place to do it. Participants use visible structure in the immediate past, including the details of syntax, as a resource for the collaborative construction of future action.

Second, through the precise way in which she pronounces "*s : : so :*" including lengthening (indicated by the colons), voice quality, and enhanced intonation (quite imperfectly represented in the transcript by bold italics), Nancy both indicates that the bit of talk now being begun has a special status of some type and displays some embodied appreciation of what she is talking about. Nancy's display of heightened appreciation thus begins before the assessment adjective itself is actually spoken. In brief, the structure of the talk in progress provides a hearer with multiple resources (including syntax and prosody) not only for hearing what has already been said but for projecting what is about to be said. The ability to make such projections is crucial because the status of the hearer is not simply that of a passive recipient but that of a co-actor, someone who can systematically produce relevant action of her own at a specific, structurally defined place. And, indeed, Tasha starts her concurrent evaluation just as Nancy's intonationally marked intensifier is coming to completion.

Turn Transition, Projection, and Human Social Organization

The simultaneous production by separate parties of reciprocal, mutually relevant action such as occurs in the data just examined provides a particularly clear example of talk as coordinated action. Nonetheless, it might be argued that concurrent assessments are rare and unusual and thus do not demonstrate ways in which participants orient to time structure that are pervasive and structurally central to the organization of talk and action in human interaction. However, work in the field of conversation analysis (which takes as its subject matter the practices that participants use to construct talk-in-interaction [see Goodwin and Heritage 1990]) abundantly demonstrates that the use of structures that provide for the prospective monitoring of movement within utterances toward a point of recognizable completion is a generic property of units used to build action within human interaction (see, e.g., Schegloff 1996). The construction of a state of talk constitutes an elementary form of human social organization, one in which action is constructed through the distributed work of socially differentiated participants. Thus turns at talk are built through the reciprocal work of both speakers and hearers (see Goodwin 1981 for analysis of how speakers systematically work to obtain the visible orientation of hearers before producing unbroken sentences), and exchanging turns in conversation requires that multiple parties change relevant social positions as the speaker becomes a hearer and someone who was a hearer becomes the new speaker. Participants are able to accomplish this exchange with fine precision. One key resource that makes this possible is the existence of *turn-constructional units* that provide a specific place, their recognizable completion, where turn transition is relevant. Hearers systematically monitor the structure of emerging talk prospectively in order to locate unit completions, upcoming moments where it will be possible for them to take the position of speaker. Time, in the form of sequential projectability, is central to the organization of such units. In their classic analysis of the turn-taking system Sacks, Schegloff, and Jefferson (1974:720) state that "whatever the units employed for the construction [of turns at talk], and whatever the theoretical language employed to describe them, they still have points of possible unit completion, points which are projectable before their occurrence." Orientation toward a richly structured prospective horizon made visible through the unit structure of language is central to the organization of talk as a primordial form of human social action.

In formal linguistics the inclusion of such a sequential framework in the basic units used to build utterances has been largely ignored. Such neglect may result in large part from the fact that linguistics has typically focused on units no larger than an individual sentence produced by an idealized, isolated speaker. Within this framework issues of social coordination and the necessity of projectability by participants who are not currently speaking do not arise. More generally, in the human sciences language has typically been analyzed almost exclusively as a symbolic system rather than as a form of social organization in its own right. For example, scholars speculating about human evolution and the origins of language frequently theorize about what social functions it might have facilitated, even suggesting that it might have been central to increases in the capacity for social coordination. However, the contributions that increasing linguistic ability might have made are described almost exclusively as an increase in representational ability. For example, Barsalou (1999:65), speculating about language as "*preparation* for situated action" (italics added), summarizes three arguments by evolutionary theorists as to how language might have facilitated "sophisticated social coordination": (1) establishing "shared beliefs about the environment,"(2) "describing actions to perform on the environment," and (3) describing "roles in groups." What such a focus on the descriptive capacity of language misses is that the production of talk-in-interaction—the primary matrix within which language emerges, is shaped, and functions as consequential action—is in and of itself a new, immensely powerful, and distinctively human form of social organization, one whose properties require detailed study (Goffman 1964). Insofar as participants engaged in talk are continuously faced with the task of coordinating situated action, from turn taking to storytelling to the constitution of hierarchy through commands and their responses (M. H. Goodwin 1990), time in the form of sequential organization is a pervasive, intrinsic component of both talk and action.

Projection in Stories

Storytelling in conversation provides an environment for investigating how a prospective orientation toward events that have not yet occurred is implicated in the constitution of a range of different kinds of phenomena, including the practices used to constitute what counts as a story. Sacks (1974) notes that stories are typically "long" turns that contain many clauses and sentences. However, producing a turn that requires multiple turn-constructional units is problematic: transition to a new speaker becomes relevant at the end of each unit. How, then, can speakers systematically produce an extended turn? One set of practices takes the form of a sequence of actions that give the story a particular shape. First, the teller produces a single unit turn, a story preface, that announces the availability of a story without actually telling it. This is followed by a reply from a potential hearer, who can either ask to hear the story or reject it. If the teller's bid is accepted, the actions of both parties have established that an extended turn is being entered, and speaker proceeds with the story while the hearer allows multiple units without attempting to take the floor at the end of each. At the completion of the story the hearer provides a relevant response of some type (as well as other structurally different kinds of comments within the story [see Goodwin 1986*b*, Schegloff 1982]).

COMPETING PROJECTIONS

Hearers are expected to provide a response at the end of the story. However, given that the story will contain multiple units, how can the hearer recognize when the story itself and not just another unit has finally come to completion? Tellers attend to this issue by providing as one component of the initial preface a characterization of what the story will contain (e.g., "The *funniest/most terrible* thing happened to me last night"). Speakers can use this characterization as an interpretive template for monitoring the unfolding story until something that would count as instance of the characterization arrives and thus locate the projected completion of the story (Sacks 1974). Such a characterization constitutes a *prospective indexical* (Goodwin 1996). Just as classical indexicals—pointing gestures, pronouns, and demonstratives such as *this* and *that*—accomplish full reference only when linked to something in their context, prospective indexicals point toward something in the future that will elaborate their meaning (e.g., what in fact constitutes the "funniest thing"). They differ from most other indexicals in that a temporally unfolding process is relevant to their structure. The recipient of a prospective indexical is put in the position of having to monitor, after the occurrence of the prospective indexical itself, further talk or other events in order to determine what precisely the term is referring to. This monitoring is structured by the interpretive template that the prospective indexical provides. Examining a specific

1 Phyllis: Mike siz there wz [a big fight] down there
2 last night.

FIG. 4. *A story preface.*

story will allow us to investigate some of the contingencies of this process.

The talk in figure 4 occurred at a backyard picnic attended by several couples, including Phyllis and Mike, Pam and Curt, and Gary and Carney. The night before, Mike had gone to a dirt-track car race. In the following Phyllis uses both past tense (*said* and *was*) and the phrase "last night" to announce the availability of a story about events that occurred prior to the present. Mike is both the source of the story and the party who actually witnessed the events that will be reported, and indeed this preface works to put him and not the speaker of the preface in the position of teller. The substance of what will be told is glossed with the prospective indexical "big fight." Faced with the task of locating the climax of the story, the place where it becomes relevant for them to produce a response to it, Mike's recipients can listen to what he reports until something that would constitute a "big fight" arrives.

However, the situation quickly becomes more complicated. For reasons that will not be examined here (see Goodwin 1986*a* for a detailed analysis of the interactive organization of this story), shortly after Mike launches the story his wife, Phyllis, provides a second, quite different characterization of what it will contain, portraying Mike's protagonists as strutting males full of empty bravado who violently threaten each other as they throw their helmets off but then "just look at each other" (line 28 in figure 5) instead of actually fighting. Other listeners associate themselves with Phyllis's new characterization. Thus Curt calls Mike's racers "Little high school kids" (line 37), and Gary says that "they know they won't get hurt" (line 35). By virtue of Phyllis's two competing characterizations of the events he is about to describe,

26 Phyllis: en den they go down
27 'n they ***thrr***ow their h***hel***mets off
28 'n then th(h)ey j's *l:l*ook at each other.
29 Mike: But, this
32 Curt: Ye: : h hh ***heh*** heh
33 Phyllis: °ehhhehhhhh
34 Mike: This: : uh : : :.
35 Gary: They know they ain't gonna get hurt,
36 Phyllis: = ehh heh!
37 Curt: Liddle ***Hig***h school kids,
...
41 Mike: De ***Wa : : :*** ld spun ou:t. 'n he ***wai***ted.

FIG. 5. *A competing characterization.*

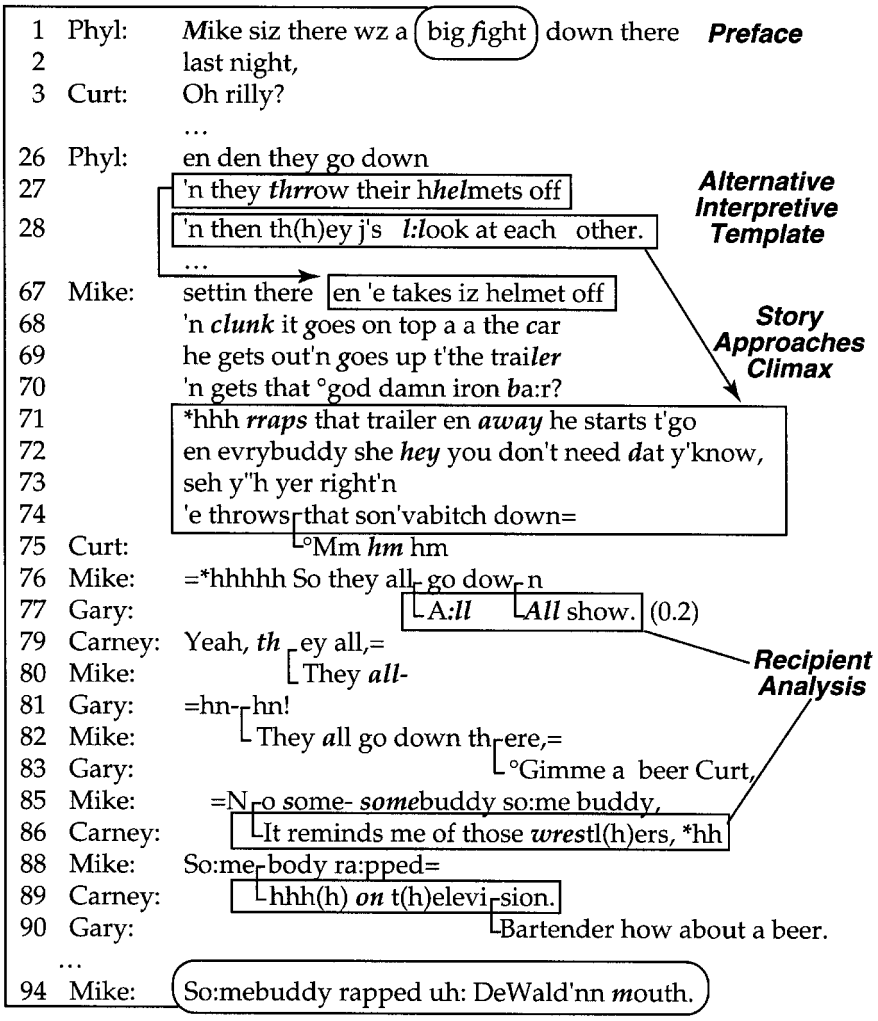


FIG. 6. *Competing climaxes.*

Mike as a teller is now in a difficult position. He is already launched into a story that is organized to be compatible with the epic battle offered in Phyllis's initial prospective indexical. However, his listeners now have in addition a second, quite different interpretive framework, one primed to locate something that can be characterized as empty show rather than an actual fight and thus to undercut the story that Mike is attempting to tell.

This competing template for monitoring the events Mike is describing creates serious problems for his telling. As the story approaches its climax, he describes events that match Phyllis's second, disparaging projection of what the story will contain (fig. 6): the principal protagonist first throws his helmet off (line 67) and then agrees to drop the iron bar with which he has been advancing toward his opponent (lines 71–74). Right after this happens, Gary, one of the recipients of the story, provides a next move to it, an analysis of the events it reports as "all show" (line 77).

The actions begun by Gary and subsequently affiliated with by Carney ("It reminds me of those wrestlers on television" [lines 86, 89]) provide both a substantive and a structural analysis of the story. Substantively what is said here undercuts the version of events that Mike is offering by treating the "big fight" as something quite different, a mere show by childish bullies who avoid actually coming to blows. Structurally, these moves, as subsequent responses to the story which provide a summary analysis of the import of the events it reported, treat the story as having come to completion. Indeed, after his "All show" Gary turns to other matters and asks for a beer (lines 83, 90). Mike, however, vigorously opposes these actions. Starting in line 80 he overlaps his audience's treatment of the story with actions that both extend it (and thus contest their analysis that it has

reached completion) and report an actual blow ("Somebody rapped DeWald in the mouth" [line 94]) and thus offer a version of events that is consistent with the original preface.

What will in fact count as the story is thus organized through a set of interlocking interactive practices being deployed by both the speaker and the hearer. A complex orientation to both temporal and sequential organization in the constitution of these units is crucial to the practices at work here (for discussion of a multiplicity of other ways in which time is relevant to the organization of stories, see Ochs and Capps 2001). As part of the interactive process of establishing what will in fact count as a story, tellers provide story recipients with interpretive templates that they use to monitor the events being reported (frequently though by no means always marked as having occurred in the past) prospectively in order to locate when the story arrives at its climax, the place where recipients are expected to provide a response to it. This quite specific orientation to time in the form of sequential organization emerges from the participants' task of collaboratively accomplishing relevant social action.

FORMULATING THE PAST TO CONSTRUCT A SOCIALLY RELEVANT FUTURE

The relevance of both formulation of the past as a warrant for current action and orientation toward future action to the construction of stories can extend well beyond the current event (fig. 7). M. H. Goodwin (1990) describes a gossip dispute activity of young girls called He-Said-She-Said. In order to bring about a future confrontation, one girl, the Instigator, uses stories to tell a second girl that a third has been talking about her behind her back. The stories, which report both what was said about the addressee in the past and how the current teller defended her there (and thus make implicit suggestions

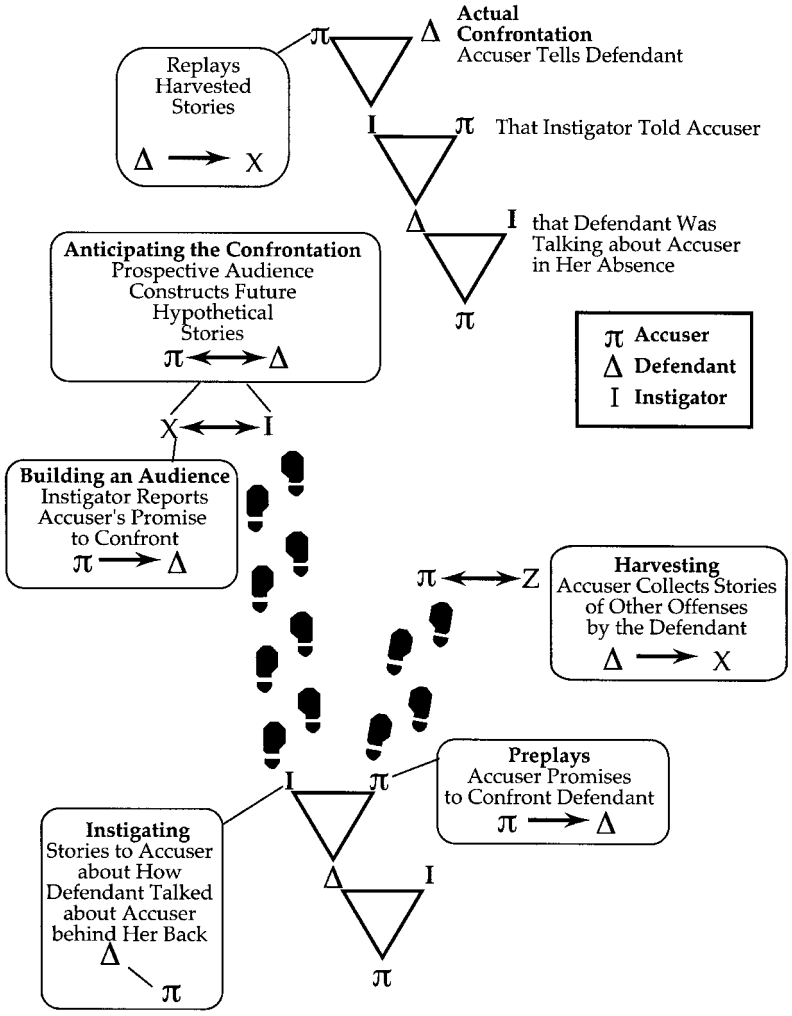


FIG. 7. *A family of stories used to organize unfolding social action.*

about politically relevant alliances), are carefully organized to elicit a reportable promise by the second girl, who will become the Accuser (π), to confront the girl who talked about her, the Defendant (Δ), when the two next meet. The projected confrontation becomes a much anticipated, newsworthy event which all of the girls in the neighborhood look forward to excitedly, a process that generates its own structurally distinct kinds of stories. First, the Instigator tells others about the promise to confront, carefully playing down her own instigation of this event. Second, she and others now produce a range of future, hypothetical stories as they project what the protagonists might say to each other and how the confrontation might go when they finally meet. Third, as part of the process of preparing for the confrontation the offended girl/Accuser harvests stories from other girls about what the offender did to them that can be used as ammunition in the confrontation to provide evidence for the flawed character of the Defendant.

This process thus generates a complex family of stories encompassing both carefully structured reports of the past and hypothetical stories about what will occur in an anticipated future. Despite the range of different kinds of temporal organization found within individual stories, the entire process is organized through a sequential orientation to the practices required to accomplish relevant future action (e.g., the Instigator reports past stories to elicit a current promise to confront someone in the future; the Accuser harvests reports of past events for use in the confrontation; the Instigator reports the promise made by the Accuser as part of the process of mobilizing not only the principal protagonists but also the audience to the future confrontation; etc.).

The data so far examined, spanning phenomena from within the clause and individual turn to extended activity systems that mobilize an entire community through families of discourse events that place structurally different kinds of participants in consequential moral positions, provide some evidence that sequential organization, which builds upon a relevant past to accomplish future action, is a pervasive component of the practices deployed by participants in interaction to build social action through talk.

Alternative Projective Environments Provided by Different Languages

Emerging syntactic structure provides participants with a key resource for projecting upcoming events within a turn and the utterance being built within it. The tightly knit syntactic structure of English, with a required subject near the beginning followed by the verb and then an optional object and prepositional or adverbial phrases, provides powerful structure for projection from at or near the very beginning of an utterance, with the effect that many aspects of its completion may be visible well before that point actually arrives. However, Fox, Hayashi and Jasperson (1996; see also Tanaka 2000) note that not all languages build utterances in this fashion and raise the very interesting possibility that such early projection may not be uniformly available. For example, in Japanese early components of an utterance project far less about what will follow them than English, and crucial information about what is being said—the verb and the stance that is being taken up, as marked with a range of postverbal particles—is provided only at the very end. There is in fact evidence that the possibilities and practices of projection in Japanese are quite different from those that occur when English is spoken and that this might account for more global observations about apparent cultural differences in interaction style, for example, that Japanese-speakers are more "involved" with each other, as is demonstrated by more extensive recipient comment within turns, etc. (Fox, Hayashi, and Jasperson 1996:212):

> If we are correct that the beginnings of TCU's [turn-constructional units] in Japanese do not provide recipient with much information about how the utterance is going to proceed, then it makes sense for speakers to produce relatively short PPU's [pause-bounded phrasal units], whose interactional implications the recipient can acknowledge or question, as the speaker works on a larger turn. . . . This allows the recipient to acknowledge small pieces without having to know exactly where the speaker is going with the full turn . . .

Moreover, it is suggested that the final particles in Japanese might have evolved precisely to provide definitive evidence of turn completion in a linguistic environment where this is only weakly facilitated by syntax (p. 213). In brief, by virtue of the way in which the syntax of different languages facilitates or inhibits the projection of relevant structure, participants using different languages may well inhabit quite distinct phenomenal worlds that have consequences in detail for how they build action together. In recent work Tanaka (2000) has provided a sustained analysis of projection in Japanese showing that it systematically differs from projection in English while also making use of resources in addition to syntax.

Such variation in the projective possibilities provided by different languages might suggest very sharp cross-linguistic differences in the possibilities for building relevant collaborative action. However, syntax is but one component of a larger ecology of sign systems used to build action within interaction (Goodwin 2000*a*, Tanaka 2000). Though sensitive to the detailed structure of emerging talk, many other practices for the construction of embodied action provide resources for the projection and organization of action that are not determined by variations in syntax. Some of these will now be briefly examined.

Complex Contextual Configurations

Much work in pragmatics and speech-act theory has taken as its point of departure Austin's (1962) question of "how to do things with words" and focused largely on structure found in the segmental structure of language (e.g., what can be displayed on the printed page through use of the alphabet—sentences and the words they contain) or on proposals about beliefs and other mental states of the speaker (unlike much work that followed his, Austin's paid significant attention to the situation in which a sentence was uttered, a christening or a wedding, for example). However, analysis of talk-in-interaction provides strong evidence that action is not built through words alone. Instead, participants accomplish action by using simultaneously quite different kinds of meaning-making practices which mutually elaborate each other. Thus an utterance typically contains both segmental structure and prosody, for example, an intonation contour. Though both of these occur within the same strip of speech, they are in fact quite different. To use terms long ago proposed by Bateson (1972), segmental structure is "digital," composed of discrete units which become meaningful through their placement in a larger set of contrasts (e.g., a small inventory of segments, such as "p" and "b," can combine with other segments to produce an enormous set of larger units, the words of a language, while distinguishing these from each other through systematic contrast, such as the way in which "p" and "b" differ in terms of voicing), while intonation and other forms of prosody are "analogic," patterns perceived as ongoing variation in a continuous whole (e.g., the rise and fall of pitch throughout an utterance). Both of these systems provide resources for making projections about upcoming events and displaying completion, but they do it in quite different ways. In the segmental stream syntax can be used to mark and recognize where an emerging unit can be seen as complete (e.g., "We put toilet paper across the door") and where it is not yet complete (in "We put toilet paper across" a noun following the preposition "across" is lacking). In the domain of intonation certain "terminal contours" can be recognized as marking different forms of completion (e.g., falling, rising, and falling rising, the latter frequently found after items in a list). Participants can recognize when a unit is not yet complete (a crucial resource for producing restarts and other visible repair initiators) and can systematically make projections about when the completion of the intonation unit will arrive (see, e.g., the discussion of interaction after "pitch peaks" in Goodwin 1986*b* and Schegloff 1996:84–86). Schegloff (1996) provides sustained analysis of the part played by both syntactic and intonational projection in the construction of turns-at-talk, and Ford and Thompson (1996) describe the importance of both syntax and intonation in the construction of "complex transition relevance places."

What consequences does the existence of more than one system for projecting unit completion have for the ability of participants to build action within interaction? Why not opt for simplicity and just use one? Syntactically a sentence can have more than one point of possible completion. For example, "I gave up smoking cigarettes" could be complete syntactically after "smoking" (e.g., "I gave up smoking") as well as after "cigarettes." Moreover, getting all the way to "cigarettes" might be consequential for the speaker (e.g., to inform his hearers that while he is no longer smoking cigarettes, he is still smoking something else—the utterance was recorded in the early 1970s). If syntax were the only system providing information about when units could be seen as complete, speakers would run the systematic risk of having another start before they had in fact completed what they wanted to say (e.g., after "smoking" when they wanted to get at least as far as "cigarettes"). However, by using a second system, such as the projections about completion provided by intonation, simultaneously, speakers can display to recipients that only a subset of the places where completion can be recognized syntactically are in fact to be treated as possible turn completions (Sacks 1995 [1992], Schegloff 1996, Ford and Thompson 1996). And, indeed, in the cigarette utterance the intonation contour over "smoking" shows that further talk will be produced before this unit reaches completion.

Though most approaches to the study of how human beings build meaning through the use of signs have focused on the analysis of single coherent systems,[2] what is relevant here is the way in which different sign systems are used together to create a whole that is greater than the sum of its parts. Moreover, what is at issue is not simply different signs but different kinds of signs, with each kind getting crucial aspects of its structure from the medium used to construct it. The continuous, changing flow of intonation permits the construction of a single pattern that can encompass multiple segments and vary in ways that permit quite subtle displays of changing stance, emotion, etc., some of which can be understood even by a person who does not understand the language being spoken. By way of contrast, building larger units from a smaller inventory of discrete linguistic segments creates the possibility of unlimited, novel combination without destruction of the units being combined. In contrast to what happens in a nonsegmental system (e.g., mixing different-colored paints together creates a new color that loses the original colors), the linguistic segments that make up larger units maintain their integrity within those units and are still visible and available for new combinations. In a spoken utterance the distinctive properties of each of these systems mutually elaborate each other (e.g., through intonation speakers can make displays about their stances toward what is being said, while segmental structure specifies what is being commented on through the intonational displays—note the assessment in figure 3). I will use the

2. For example, while recognizing the possibility of a general science of signs which he called *semiotics*, Saussure (1966) and those who followed him found it most profitable to focus on a single system such as language.

term *semiotic field* to refer to a system of signs and the medium used to build those signs.

In face-to-face interaction many different kinds of semiotic fields can contribute to the organization of action. In addition to intonation and linguistically meaningful segmental structure, utterances can be accompanied by gestures of different types, and sequences of talk can be framed by participation frameworks made visible through the mutual orientation of the participants' bodies. Moreover, as we have seen in the work of the archaeologists with their forms and Munsell chart, material structure in the setting where talk is occurring can contribute to its organization in a variety of ways (Goodwin 2000*a*). The particular set of semiotic fields that the participants treat as relevant to the organization of action at a particular moment I will call a *contextual configuration*. Contextual configurations can change both within an encounter (indeed, within a single utterance [see Goodwin 2000*a*]) and systematically from setting to setting (e.g., telephone calls include only a subset [but a most crucial subset] of the semiotic fields available to participants who are physically co-present). Each of the semiotic fields in a contextual configuration can provide different kinds of information about the temporal and sequential organization of both the actions in progress and the state of the current interaction. The effect of this is that participants in interaction are using a complex ensemble of structurally different kinds of practices to organize the time structure of both the events being reported through talk and the activities implicated in the construction of a state of talk.

Embodied Time in Storytelling

Some of these phenomena can be examined in the brief strip of conversation in figure 8. Nancy and Tasha had both attended the same two-year women's college, though at different times. Here Nancy is describing pranks they played on someone in her dorm who was disliked. That person has already been introduced. Figure 8 contains two brief, linked subsequent stories, one about putting Vaseline on the disliked dorm-mate's doorknob and the next about covering the door to her room with a network of toilet paper. Because of their status as "second stories" (Sacks 1995 [1992]) no preface is necessary and the descriptions are brief. The story about the toilet paper contains both a background segment setting the stage for the climax (putting the toilet paper on her door [lines 5 and 6]), the climax itself (her having to walk through the toilet paper [lines 7–10]), and subsequent responses by both parties (lines 11 and 12). The use of tense and aspect in the grammatical organization of the report is interesting and relevant but will not be further examined here.

PARTICIPATION FRAMEWORKS

Kendon (1990) has described how encounters in face-to-face interaction are made visible and sustained through the mutual orientation of the participants' bodies, what he calls an "F formation." In figure 8 the lower bodies of the participants are aligned toward each other, and this alignment is sustained throughout the recording (approximately 11 minutes). This mutual alignment constitutes one component of a *participation framework* that marks the temporal duration of extended strips of discourse in face-to-face interaction. Moreover, though this will not be examined here, it has sequential relevance as well. Changes in posture provide participants with resources to display their availability for entry into an encounter and to negotiate exit from it.

Within this overall framework of sustained mutual availability, participants can mark alternative states of engagement and disengagement (Goodwin 1981) through phenomena such as gaze toward or away from their coparticipants. In the midst of the telling in figure 8 both parties are gazing intently toward each other and thus displaying heightened mutual involvement, a positioning appropriate to the collaborative production of talk by a speaker for a hearer. However, at the beginning and the end of the sequence they are looking away from each other, without, however, dismantling the framework of mutual orientation visible in their lower bodies (fig. 9). This positioning is quite appropriate to parties who are co-present but not talking to each other at the moment. Through the combined use of posture and gaze these participants thus create embodied frameworks that display and visibly sustain local spates of focused interaction intermixed with periods of disengagement. Both engagement and disengagement occur within a larger pattern of participation in an ongoing encounter that is marked by the continuing orientation of the participants' lower bodies toward each other. Different parts of the body thus create different temporally relevant displays about different components of their current interaction—on one level the encounter as a whole and on another spates of focused interaction and talk within it.

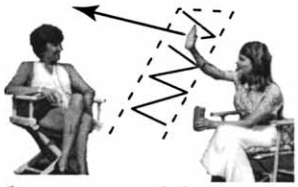
GESTURE

Within the story itself (lines 5–8) the speaker uses her hand and arm to make a series of gestures. Structurally, gesture is quite different from the participation displays being made through posture. Temporally, the gestures have a much shorter duration. Moreover, as signs they are linked to the content of what is being said. By way of contrast, the participation displays help constitute the state of mutual orientation between speaker and hearer required for the production of whatever talk will occur, without, however, dealing with any specific content. These different forms of embodied activity are implicated in quite different aspects of the ensemble of practices required to construct a telling. However, rather than being entirely separate, these alternative semiotic fields work together to create a larger complex of action. Thus, like the talk, the gesture is framed by the participants' visible mutual orientation and indeed requires it to function as a communicative action. Note, for example, how the speaker's hand is positioned right in her

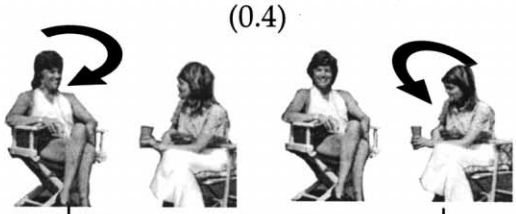
1 Nancy: Tch! ***We*** couldn't sta:nd her. so ba:dly
2 we uetuh ***t***ake her doorknob'n
3 grease it wi(h)th va(h)selin(h)ne(h)e(h)v(h)***ry n(h)***i(h)ight,
4 *HHSSSS! O(h)r,

5 *hhh y'know (if) sh'd be in th'room'n we'd-

6 *hh kinda put ***toi***let paper a***c***ross
7 so thet when she open'up the door



8 she'd aftuh co:me, ***yih*** kno:w.
9 (0.4)



10 Nancy: la ┌:m***b***astin ┌g through it
11 Tasha: └°Oh::: └°***G***hho : : d]=
12 Nancy: =***O***h: : God it wz r'lly funny.
13 (0.7)

FIG. 8. *Embodied time in storytelling.*

addressee's line of sight. The framework required to analyze this gesture is one that encompasses the bodies of both participants (i.e., the participation display) as well as the talk in progress (not all gestures are organized to be seen as communicative acts in this fashion).

As an activity in its own right a gesture contains its own temporal and sequential organization. Moreover, though occurring in quite different semiotic fields, talk and gesture are parasitic on each other—in these data each requires the other to be properly understood. Examination of this process will begin with two observations about the sequence in figure 8: (1) Lines 6–7 are

FIG. 9. *Shifting engagement.*

grammatically anomalous—no noun completes the prepositional phrase begun with "across" in line 6, though the required noun seems to be "door," which occurs later as part of another syntactic construction in line 7. (2) The climax of the story seems to occur in line 10, when the person in the room opens the door and moves through the toilet paper ("lambasting through it"). However, rather than marking that place as a story climax through heightened involvement, the participants seem to treat the story as having already reached completion by withdrawing their gaze from each other there.

During line 6 Nancy's moving hand draws an extended zigzag shape in front of her. It clearly seems to be gesture, but what does it represent? In conversation participants are regularly able immediately, indeed almost transparently, to see the meaning in the gestures of a speaker's waving hand. This is, however, very much a situated accomplishment and, moreover, one that has to take into account semiotic fields other than the hand itself. Gesture and the talk that co-occurs with it mutually elaborate each other. The problems posed when relevant talk is lacking are dramatically illustrated in the case of a man with severe aphasia who makes extensive use of gesture but can produce almost no talk (Goodwin 2000*b*). It takes considerable work by his recipients to work out what he might be gesturing about. In these data Nancy first, in line 5, uses a simple gesture to link the gesture space in front of her to the "room" she is talking about. The zigzag hand which follows not only occurs in a space that has just been marked as pertaining to the room but is performed during "kinda put *toi*let paper across." Through such temporal binding the waving hand can be seen as enacting the activity described in the talk. Moreover, at the beginning of the gesture the speaker looks above her head, begins the gesture there, and then has the ends of each line in the zigzag mark a larger rectangle. By virtue of both the co-occurring talk, which has established as a frame the room and its standard parts, including a rectangular door taller than the people who walk through it, and the precise structure of the gesture, the enactment could be seen as depicting actions performed on the door of the room. Moreover, by not explicitly saying "door" in a syntactic slot where it would be required if talk were the only semiotic field in play, the speaker displays that she is expecting her addressee to be able to see the unsaid, that is, to take into account the way in which talk and gesture mutually inform each other within a larger whole.

PROJECTING THE STORY CLIMAX

By virtue of the way in which the protagonists, the scene, and the complicating action are established by the end of line 7, telling the climax as news or "new information" can be problematic. By this point in the story a hearer knows what the climax of the story will be: when the woman in the room opens her door (the situation described in line 7) she will have to walk through the web of toilet paper. Indeed, the recipient is nodding from the start of line 7. The teller systematically displays that she recognizes that her addressee is performing such projective work. Rather than continuing with the description, at the end of line 8 the speaker explicitly marks that her audience can already see what is about to happen by saying "*yih* know." As she says this she depicts the climatic scene visually by bringing her hand back and then thrusting it through the space where she has just enacted the door with its web of toilet paper. Here, rather than line 10, is where the climax of the story actually occurs. Its embodied enactment both takes into account the active projective work of her audience (the speaker doesn't tell her recipient something that she can be expected to know already by providing a simple description) and invites enhanced participation in the appreciation of the event (see M. H. Goodwin 1980). The description eventually provided in line 10 is a subsequent gloss of the event the audience has just seen rather than a climactic story clause reporting new information (note, for example, the metaphoric "lambasting").

The gesture that Nancy uses to enact what happens has a complex narrative and sequential organization of its own. Thus, the hand thrusting through the toilet paper requires for its proper understanding that the viewer also see as still present, as a kind of ghost image still inhabiting the space between the participants, the prior hand movements depicting the toilet paper being laced across the door, and these in turn build upon the earlier gesture linking this space to the "room."

It is common for analysts to lump together, as "nonverbal behavior," everything that participants in inter-

action do that is not linguistic. Here we see how such a simple dichotomy falls apart. The visible body is constructing through gesture quite different kinds of temporal organization and reference from the simultaneous participation frameworks being organized through posture and gaze.

STRATEGIC DISENGAGEMENT

The postural configurations that frame extended strips of talk might seem to lack the dynamic organization of gesture. However, they too are subject to ongoing interactive negotiation as participants formulate for each other the status of their current interaction. As Nancy says "lambasting," Tasha looks away from her. One syllable later Nancy demonstrates how she has understood this move and its consequences for the continued viability of her telling by also looking away. Nancy's story sequence then ends with a summary assessment (line 12) framed by postures of mutual disengagement. A moment later Tasha initiates a new telling of her own. It is at least plausible that Tasha, instead of using gaze to dismantle this facing formation in preparation for establishing a new one in which she is the teller, could have continued to display engagement with Nancy's line of talk. It can also be noted that Tasha's move into disengagement juxtaposes a number of different semiotic fields—gaze withdrawal, an appreciative assessment, and lowered volume—to build a complex display that combines appreciation of what has just been said with systematic withdrawal from it.

MULTIMODAL TEMPORALITIES

The story that occurs here is interactively organized through a multimodal contextual configuration in which the participants orient to a range of quite different kinds of time organization instantiated in diverse media. The content of the story—what is being reported—is produced through a mutually elaborating combination of talk and gesture. Indeed, the speaker explicitly displays that she is expecting her addressee to take gesture into account by not producing elements of talk that are otherwise required syntactically. Both the gesture and the talk incorporate distinctive forms of temporal and sequential organization which allow the hearer to project events that have not yet been described. Thus, to see a gesture as a meaningful sign (indeed, the climactic scene in the story) a hearer must first use the talk that accompanies it to find a relevant sense for the speaker's waving arm and then synthesize into a larger whole a succession of different hand movements that appear and disappear through time. These sign systems of talk and gesture are framed by a range of other time-relevant displays which deal with the constitution of the state of talk itself rather than what is said within it. With their lower and upper torsos, heads, and gaze the participants display to each other hierarchical frameworks of mutual orientation to the encounter as whole and local spates of engagement and disengagement within it. All of these signs are organized as relevant action through the interactive organization of mutual accountability. Thus, rather than simply producing signs, the participants treat their coparticipants as reflexive actors. They expect each other to take into account for the organization of subsequent action the projective frameworks provided by different sign systems, for example, to see the climax of the story in a gesture that builds on projections provided by earlier talk and gesture or to treat gaze withdrawal as incipient disengagement.

It is common at present to talk about multimodality in terms of comparatively recent technology, such as movies and the computer. However, data such as those just examined suggest that dense contextual configurations built through quite different kinds of signs instantiated in talk and the body constitute a primordial multimodal environment for the constitution of human meaning, knowledge, and action. Nancy expects her addressee to understand her story by synthesizing the projections provided by different kinds of sign systems working together in a rich multimodal performance. Moreover, as the work of the archaeologists with the Munsell chart demonstrates, these embodied frameworks for the production of human action can incorporate diverse semiotic structure embodied in tools and artifacts as well.

The importance of the embodied displays seen here has methodological implications for the analysis of discourse units such as stories and for the way records are made of such events. Frequently the data used for the study of stories consist simply of either a written text or a transcription of an extended bit of talk by a single speaker (or frequently just excerpts from such a unit) (note, for example, the myths studied by anthropologists such as Lévi-Strauss). The phenomena being examined here reveal the importance of having data that provide a record of not only the talk of the principal speaker but also the visible actions of the story's addressees, hearers, and other kinds of recipients. If the record of the stories in figure 8 included only the talk produced, then instead of emerging as a complex embodied speaker engaged in ongoing analysis of the relevant projections being made by her addressee, Nancy would be depicted as someone who didn't speak in fully grammatical sentences.

Conclusion

What is striking about the ways in which time is implicated in the organization of action within human interaction is the range of different kinds of temporal frameworks that participants orient to simultaneously in building even simple, brief actions. For example, as seen in the talk in figure 3, linguistic structure provides resources that can be used simultaneously to (1) structure time in the world being represented through the talk, for example, situate events being reported in a past, and (2) provide hearers with resources for projecting future events in the current interaction, the world of the telling. Similarly, in scrutinizing a tiny bit of dirt the

archaeologists in figures 1 and 2 were describing a past world being investigated (a Native American settlement) through practices that used tools from a professional past (the coding sheet of the senior investigator, the history of research encapsulated in the Munsell chart, etc.) to build a work-relevant future (the records that would form the basis for subsequent analysis).

Central to all of this is that different kinds of time are constructed and made relevant through public practices built from materials instantiated in diverse media within ongoing processes of reflexive human interaction. Some features of this process deserve more comment. Consider first its *public* character. Since the cognitive revolution of the 1960s, great emphasis has been placed upon analyzing both language and cognition as thoroughly mental and largely private processes. Thus, linguistics, which was part of anthropology at the beginning of the 20th century, now has its strongest ties to psychology, and formal linguistics is programmatically uninterested in how social phenomena might be relevant to language structure. However, as the data examined in figures 3–8 demonstrate, parties other than the speaker use emerging language structure to make inferences about upcoming events so as to accomplish collaborative action. The organization of language structure extends beyond the mental life of an isolated speaker as it engages the minds of coparticipants in distinct practices of inference and action tied to the organization of the talk in progress (e.g., acting at the climax of a story). Moreover, the use of language structure to build multiparty action through talk-in-interaction constitutes a primordial and perhaps the most pervasive site for the constitution of distinctively human forms of elementary social organization.

Second, such focus on the public organization of interactive practices provides an expanded picture of human mental and cognitive life. Rather than simply producing language and other semiotic structure, participants in interaction are attributing complex cognitive and inferential practices to their coparticipants and taking these into account in the detailed organization of ongoing social action. Thus, as was seen in the data examined in figures 4–6, tellers provide their hearers with prospective indexicals to organize their monitoring and coparticipation in an unfolding story, and these in turn provide hearers with resources for both negotiating and disputing the import and structure of the emerging story. Such reflexive practices of mutual monitoring, in which participants treat each other as active cognitive and social actors, encompass not only structure in the stream of speech but also the visible body as a public locus of the constitution of temporally unfolding meaning and action (e.g., the data in figure 8). The interactively organized sequences within which such projections and inferences are used to build action through time provide both a matrix for the organization of such projection and inference and a proof procedure for publicly testing their success.

Third, time is an intrinsic component of the *units* used to build events and action within situated human interaction. The location of relevant units is a key task facing both analysts (e.g., What are the units used to construct utterances and sentences? What are the parts of a narrative? etc.) and participants (e.g., Where in the stream of speech can speaker change occur? Where in a story are responses of different types appropriate? etc.). It might seem analytically useful first to describe units without reference to time and then to treat orientation toward time when using units to build action as a separate, distinct phenomenon (and, indeed, this is precisely the route chosen by many research programs). However, as the data examined in figures 3–8 demonstrate, aspects of a unit's structure as basic as its boundaries are constructed from their start with an orientation toward the time-relevant analysis in which others using the unit as a framework for the organization of action will be engaged. Thus storytellers begin by providing their listeners with prospective indexicals so that hearers will be able to recognize when the story comes to completion. Where the end of a unit, such as a story, is located can be negotiated and contested by different participants in ways that are consequential to its organization (see the data in figures 6 and 8).

The specific media used to instantiate particular semiotic systems provide participants with crucial resources for building relevant temporal frameworks. The holes in a Munsell chart (figs. 1 and 2) create an environment for the production of a new action within a graphic field encapsulating a history of past research. The use of multiple media to build action provides participants with crucial ability to modify the boundary projections of any single system. Thus a prosodic contour can override points of possible unit completion that are located by grammatical organization. In short, orientation toward diverse forms of time organization is built into the units and tools used to construct human action.

Fourth, the time frameworks relevant to the organization of human action encompass not only language structure and the body but also material artifacts, documents that link the current encounter to other settings, projected political disputes (e.g., fig. 7), historically shaped tools, the organized work practices of a social group, etc. The processes through which these phenomena are used to build action within situated human interaction provide one place where the details of language use, the constitution of meaning through culture, pervasive and elementary forms of human social organization, and the historical structuring of the material world by social groups can be investigated as components of a single integrated process. Indeed, it would seem that something like the ecology of sign systems which has developed around the bodies of multiple participants using talk and tools to perform temporally unfolding action in a consequential surround constitutes a crucial environment for the evolution of human language, cognition, and social organization.

References Cited

AUSTIN, JOHN LANGSHAW. 1962. 2d edition. *How to do things with words*. Oxford: Oxford University Press.

BAKHTIN, MIKHAIL. 1981. *The dialogic imagination: Four essays*. Edited by M. Holquist, translated by C. Emerson and M. Holquist. Austin: University of Texas Press.

BARSALOU, LAWRENCE W. 1999. Language comprehension: Archival memory or preparation for situated action? *Discourse Processes* 28(1):61–80.

BATESON, GREGORY. 1972. *Steps to an ecology of mind*. New York: Ballantine Books.

BERLIN, BRENT, AND PAUL KAY. 1967. *Universality and evolution of basic color terms*. Laboratory for Language Behavior Research Working Paper 1.

———. 1969. *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

BRUNER, JEROME, R. R. OLIVER, AND PATRICIA M. GREENFIELD. 1966. *Studies in cognitive growth*. New York: John Wiley.

FORD, CECILIA E., AND SANDRA A. THOMPSON. 1996. "Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns," in *Interaction and grammar*. Edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, pp. 134–84. Cambridge: Cambridge University Press.

FOX, BARBARA A., MAKOTO HAYASHI, AND ROBERT JASPERSON. 1996. "Resources and repair: A cross-linguistic study," in *Interaction and grammar*. Edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, pp. 185–237. Cambridge: Cambridge University Press.

GOFFMAN, ERVING. 1964. "The neglected situation," in *The ethnography of communication*. Edited by John J. Gumperz and Dell Hymes, pp. 133–36. American Anthropologist 66(6), pt. 2.

GOODWIN, CHARLES. 1981. *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

———. 1986*a*. Audience diversity, participation, and interpretation. *Text* 6:283–316.

———. 1986*b*. Between and within: Alternative treatments of continuers and assessments. *Human Studies* 9:205–17.

———. 1996. "Transparent vision," in *Interaction and grammar*. Edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra Thompson, pp. 370–404. Cambridge: Cambridge University Press.

———. 2000*a*. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32:1489–1522.

———. 2000*b*. "Gesture, aphasia, and interaction," in *Language and gesture*. Edited by David McNeill, pp. 84–98. Cambridge: Cambridge University Press.

———. 2000*c*. Practices of color classification. *Mind, Culture, and Activity* 7(1 and 2):19–36.

GOODWIN, CHARLES, AND JOHN HERITAGE. 1990. Conversation analysis. *Annual Reviews of Anthropology* 19: 283–307.

GOODWIN, MARJORIE HARNESS. 1980. Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry* 50:303–17.

———. 1990. *He-said-she-said: Talk as social organization among black children*. Bloomington: Indiana University Press.

GREENFIELD, PATRICIA MARKS, AND JEROME S. BRUNER. 1966. Culture and cognitive growth. *International Journal of Psychology* 1:89–107.

JEFFERSON, GAIL. 1973. A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica* 9:47–96.

KENDON, ADAM. 1990. "Spatial organization in social encounters: The F-formation system," in *Conducting interaction: Patterns of behavior in focused encounters*. Edited by Adam Kendon, pp. 209–38. Cambridge: Cambridge University Press.

OCHS, ELINOR. 1979. "Transcription as theory," in *Developmental pragmatics*. Edited by Elinor Ochs and Bambi B. Schieffelin, pp. 43–72. New York: Academic Press.

OCHS, ELINOR, AND LISA CAPPS. *Living narrative: Creating lives in everyday storytelling*. Cambridge: Harvard University Press.

SACKS, HARVEY. 1974. "An analysis of the course of a joke's telling in conversation," in *Explorations in the ethnography of speaking*. Edited by Richard Bauman and Joel Sherzer, pp. 337–53. Cambridge: Cambridge University Press.

———. 1995 (1992). *Lectures on conversation*. 2 vols. Edited by Gail Jefferson, with an introduction by Emanuel A. Schegloff. Oxford: Basil 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.

SAUSSURE, FERDINAND DE. 1966. *Course in general linguistics*. Edited by Charles Bally and Albert Sechehaye in collaboration with Albert Riedlinger. Translated and with an introduction by Wade Baskin. New York: McGraw-Hill.

SCHEGLOFF, EMANUEL A. 1982. "Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences," in *Georgetown University Roundtable on Languages and Linguistics*. Edited by D. Tannen, pp. 71–93. Washington, D.C.: Georgetown University Press.

———. 1996. "Turn organization as direction for inquiry into grammar and interaction," in *Interaction and grammar*. Edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra Thompson, pp. 52–133. Cambridge: Cambridge University Press.

TANAKA, HIROKO. 2000. Turn-projection in Japanese talk-in-interaction. *Research on Language and Social Interaction* 33(1):1–38.