

# Introduction: Human Sociality as a New Interdisciplinary Field

*N. J. Enfield and Stephen C. Levinson*

At the heart of the uniquely human way of life is our peculiarly intense, mentally mediated, and highly structured way of interacting with one another. This rests on participation in a common mental world, a world in which we have detailed expectations about each other's behavior, beliefs about what we share and do not share in the way of knowledge, intentions, and motivations. That itself relies both on communication (linguistic and otherwise) and on a level of cooperation unique in the animal world. This mode of cooperative, mentally mediated interaction enables the accumulation of cultural capital and historical emergence of cultures. By inheriting a world of social organizations and values, individuals are released from reinventing the wheel. In turn, cultural capital shapes the style of interaction in local social groups, hiding shared commonalities behind the veil of distinct languages, cultural styles, and forms of social organization.

This, at least, is the thesis of this book. It brings together anthropologists, linguists, psychologists, and sociologists whose work has not been juxtaposed before. When we put the pieces of the jigsaw together, what emerges is a new map of a still underexplored terrain—the roots or foundations of human sociality.<sup>1</sup> We propose that this is a new scientific domain, a coherent subject for investigation constituted by intersecting principles of different orders (ethological, psychological, sociological, and cultural) that work together to produce an emergent system, a system of human sociality and social interaction.

In this introductory chapter, we want to give readers a sense of how the rest of the chapters fit together to form an outline of this domain.

We first sketch some contributing research traditions and the ways they fit together. We go on to delineate the different phenomena that are the focus of the individual chapters, drawing attention to the connections that run through them. Finally, we sketch our own synthesis of the domain.

The ideas in this book ramify and connect with one another in multiple ways. Although no linear order of chapters could capture such a network of connections, our division of the book into five parts aims to emphasize certain linking themes. Part 1 consists of four chapters focusing on central properties of face-to-face interaction, the arena in which human sociality is centrally exercised. Part 2 focuses on psychological foundations of human sociality, exploring the question of just what it takes to pull off human interaction as we know it. Part 3 deals with issues of culture and cultural difference, and the ways sociocultural forces may play a role in structuring interaction and interactional expectations, and vice versa. Part 4 explores ways in which cognition is defined by its being exercised in social interaction, and how the social exercising of cognition has effects both on our understanding of the individual's psychology (part 2) and of the higher levels of social organization and broader cultural conventions (part 3). Part 5 features phylogenetic perspectives, with two chapters asking how key features of the human system for interaction could have evolved and a third chapter comparing human social abilities with those of the other great apes.

### **Distinctive Properties of Human Sociality**

The focus of this book is the distinctive nature of human sociality, the character of the social interaction that underpins social life. We do not mean its mere complexity. Many animals, not just humans, have complex social lives. Ants, for example, have hierarchies, complex divisions of labor, advanced fungal agriculture, communication, organized transport, colonization, and warfare. In this chemical society, the essential glue that holds the vast ant communities together is pheromonal. Our inquiry into the roots of human sociality asks about the nature of the special kinds of social bonds that set humans apart. Our brand of sociality distinguishes us even from our nearest relatives, the apes. We perhaps share with the apes some basic social principles: flexible coalitions, out-marriage, short-lived hierarchies. But these few commonalities are not going to explain the divergences: human advanced agriculture, elaborate communication systems, organized transport, planned colonization or

warfare, to mention a few antlike properties. Above all, the primate background does not explain the extraordinary cultural variety of human social organization, communication, and lifestyle. The entire enterprise of ethnographic research has been dedicated to understanding this diversity, and the rich details of cultural worlds that individuals inhabit. Less attention, certainly by anthropologists, has been paid to understanding the commonalities, the shared foundations in human cognition, motivation, instinct and social interaction that make these variations possible. Here, we know much less because standard social inquiry trades on these commonalities (e.g., in participant-observation) without examining them, being prompted mostly by the discovery of difference. But there is a hidden raft of commonality that makes the expression of difference possible in the first place.

Supported by uniquely human abilities, and responsive to context-specific motivations and accumulated cultural conventions, human social interaction exhibits striking properties not found elsewhere in the animal world. It involves frequent, intense, and highly structured interaction, using complex communication systems, on which the rest of culture depends for its realization. Robust parallels across cultures in the organization of everyday talk suggest an ethological foundation to human interaction. But above all what makes human interaction qualitatively distinct in the animal kingdom is that it is built on intersubjectivity, enabling a brand of joint action that is truly open-ended in goals and structure. This provides the building blocks for human cultural diversity.

Uniquely human phenomena such as cooperation, commensality, morality, and the inhibitions that underlie it, prolonged dependence of offspring, capacity for intention attribution, planned deception, and the highly structured nature of social interaction form an interdependent network. The researcher may be positioned at any point on this network and see human sociality as branching from that point. One might say, for example, that cooperation is the key: It is cooperation that makes morality essential; allows collective investment in offspring; and lies behind the sheer interest in social interaction, the special communicative abilities, and the cultural shaping of shared lifestyle. Other starting points are possible. Different authors in this book start from different corners of this network (see also Kockelman 2005), but, crucially, they agree that human social life is intricately structured through the attribution of actions, motives, intentions, and beliefs to fellow interactants. (They do not always agree, however, as to the best analysis of how such attribution is achieved; see below.)

## Contributing Research Traditions

Ideas presented in this book relate historically to a number of major strands of research that have hitherto remained largely unconnected. One of our motivations as conveners of this project was to bring different traditions together, and allow a common focus on the foundations of human interaction to emerge. We outline four domains: Theory of Mind (ToM), Gricean pragmatics, the analysis of talk and action in interaction, and related developments in anthropology.

### *ToM and the Psychology of Human Interaction*

In developmental psychology, a keen research interest has arisen in the human ability to attribute knowledge, intentions and beliefs to other humans, and to monitor these attributed inner states, using such ongoing models to interpret actions and events.<sup>2</sup> Some developmental disorders such as autism can be understood as a failure to achieve this level of understanding (Baron-Cohen 1995).

The development of an understanding of other minds in human children takes a partly puzzling course. A comprehensive understanding of others' inner states, and especially that others might have false beliefs, develops at around four years of age, surprisingly late in childhood (see Astington for discussion and references). By comparison, normal infants' mental mastery of the material world (naive physics) seems complete by even one year of age, and the fundamental grammatical structures of a language are well in place by three at the latest. Some theorists suggest that children gradually have to construct a fully articulated set of psychological skills for modeling and reasoning about others' internal states—that is, what others might want, think, feel, and know (or not know, as the case may be). This set of skills is known as ToM (in the sense of an actor's "theory," not of an analyst's).

An apparently difficult and particularly late-developing component of ToM is the ability to attribute false belief. For example, a child in the course of development learns that whereas he or she knows that chocolate in a chocolate box has been replaced by pencils, others may nevertheless expect there to be chocolate there. Children under the age of four do not show evidence of being capable of attributing false beliefs like this. (These false-belief tasks are sometimes taken to define ToM in a strict sense.<sup>3</sup>)

But here is the puzzle. It seems on first principles that without beliefs about what others do or do not know one could not be a competent interactant. How would you know what to tell me or what would require

pointing out? How and why would you even begin communicating if you had no developed concept of other minds? From age one and even earlier, children show much evidence of taking other's beliefs and intentions into account. Their development rests on it (see Gergely and Csibra, Liszkowski, Tomasello). That children of 12 months use pointing gestures to inform—for example, telling an adult the location of something the adult is looking for—shows that they have the ability not only to produce action oriented to another's mental state (e.g., someone not knowing where something is) but to presuppose that the action (e.g., a pointing gesture) will be recognized by another as having a communicative intention. There is much already in place by age one.

Clearly, then, ToM is a matter of incremental mastery. Among the issues this volume struggles with are what the components of a ToM must be, what the incremental stages are, and at what point down that incremental hierarchy we share elements of ToM with our nearest cousins among primates. Judging from ape behavior, it seems harder to understand beliefs than desires. Judging from human infants, it is harder to understand second-order beliefs (John thinks that Mary believes the chocolate is in the box) than first-order ones (Mary believes the chocolate is in the box). Does effective joint action—a possible precursor to culture—presuppose ToM? Is language crucial for discovering the full potential of a ToM? Are there culture-specific practices that encourage, or constrain, its development in childhood? With a better grasp of these components, the stages by which they develop, and the ways they are deployed in daily life, we may better understand both human ontogeny and phylogeny.

### *Gricean Pragmatics*

A second line of work guiding the debates in this book originates in philosophy, specifically in H. P. Grice's (1957) idea that meaning is grounded in the recognition of intention. Seeing you fall over ahead of me up a steep path, I am relieved to see you get up and wave in my direction, taking your wave as designed to make me think you are OK. The wave works because you have correctly calculated that I will recognize the plan behind your action, namely getting me to recognize that you intend me to think you are OK. In this example, the wave has a nonce or one-off meaning recoverable against a background of your figuring what I would figure when I see it.

Grice's idea is important because it shows meaning, in a broad sense, to be independent of language or convention. This points to possible precursors to conventional meaning, in ontogeny, diachrony, and,

perhaps, phylogeny. On this account, meaning is not a property of signs or symbols, but a property of minds in (mediated) interaction with other minds. Conventional meanings can be thought of as arising from repeated use of what were once novel signals. If I fall down and likewise wave, we might set up a miniconvention that then spreads through the community of hikers.

Another important aspect of this psychologizing of meaning is that it allows us to analyze the unspoken communicative contents associated with conventional symbols. For conventional meanings never exhaust the import of what is said. The simplest utterance usually carries with it a penumbra of intended but unspoken thoughts. (Consider *What are you doing tonight?* which is likely to be forecasting an invitation, not simply asking a question.)

The whole business of exchanging intentions in communication relies on background assumptions that help to narrow the range of intention attribution. Grice (1975) suggested that the essential background assumption by which interactants constrain and guide their inferences about speaker intentions is a principle of cooperation. (The principle, comprising maxims of quality, quantity, relevance, and manner, has since been updated in modern recastings such as Levinson 2000 and Sperber and Wilson 1986.) Recipients of others' signals work on the assumption that such signals have been designed specifically for them to extract the intended meaning. In turn, senders of such signals design those signals in such a way as to take into account such an expectation of targeted design on the part of hearers. By a principle of audience design (or "recipient design"; Sacks and Schegloff 1979), any utterance should have been formulated by a speaker with the intention that it cause just the right effect in the receiver, taking into account the common ground of the particular combination of speaker and addressee(s). For example, in telling you something about my colleague John, I will first refer to him in a way appropriate to your knowledge of him—for example, as *John* if we commonly know him as John, but as, say, *a colleague of mine* if I suppose you have never met him (Enfield and Stivers in press).

In sum, Gricean principles require the modeling of others' inner states, and thus presuppose a ToM. They also entail a stock of common ground, readily provided by culture (e.g., that *How do you do?* is not seeking information, that it is OK to strip down on the beach but not on the street, or that sweet desserts come after savory main courses; Enfield 2000; Levinson 1995). E. Goody (1995) suggests that the entire structure of social roles in a society should be understood against this background, providing systematic constraints on appropriate social intentions and their ascription.

### *Microanalysis of Social Interaction*

Detailed study of the systematics of social interaction in its own right was initiated by a string of 20th-century mavericks including G. Bateson, R. Birdwhistell, H. Garfinkel, and E. Goffman. The study of the systematics of social interaction has since passed largely to conversation analysts and other students of talk and action in interaction, with research resulting in a detailed inventory of observed interactional practices and patterns.<sup>4</sup> Most of these practices can be characterized as sequences of interlocking social actions (e.g., turns at talk) whose interpretations are associated with specific (sometimes culture-specific) expectations and preferences. The taking of turns at talk, the openings and closings of conversations, the structure of request sequences, practices for correcting or repairing utterances, and so on, have been carefully explored in English-language conversation. There is also an increasing knowledge of how these things work in other languages. Emerging from this research are candidate universals for the organization of human interaction (see Schegloff), such as the mechanism for transition of turns at talk in informal conversation, and the ways in which interlocutors correct and repair their own and others' utterances—a crucial mechanism for maintaining intersubjectivity. There is a strong expectation that such structures should be universal, given that they are essential for preserving order and agreement in moment by moment social experience. (See Goffman 1981:14, for a list of such "system requirements and system constraints," including "framing capabilities," Gricean principles, and "nonparticipant constraints.")

Conversation analysts try to avoid the psychological turn that characterizes ToM research and Gricean pragmatics. They prefer to talk in terms of actions as recognizable through the details of their observable structure and their specific placement in sequences of action. But they are equally interested in intersubjectivity, the way in which a shared understanding is arrived at. Hence the special interest in "intercalibrative" mechanisms like repair and audience design.<sup>5</sup>

### *Related Developments in Anthropology*

Biological anthropology has entertained various solutions to its central puzzle of the evolution of human cognition and language, conceived in terms of causal adaptational pressures (e.g., Dunbar et al. 1999; Richerson and Boyd 2004). Most if not all of these hypotheses have an interactional flavor. Primatologists and comparative psychologists have tried to pinpoint exactly what properties humans share with their nearest

primate relatives, and they have tried to probe to what extent deception, and more generally ToM capabilities, extend across the higher primates (Byrne and Whiten 1988; Sussman and Chapman 2004; Tomasello et al. 2005; Whiten and Byrne 1997). Just as the ability for planned deception has been seen as a defining Rubicon in human evolution, so imitation has been seen as an important threshold for the possibility of acquiring culture. The cooperative nature of human interaction raises fundamental evolutionary puzzles (Boyd and Richerson 2005, this volume; Henrich et al. 2004). Cooperative instincts are unlikely to evolve and persist under natural selection, which means there must be higher-level checks on good intentions, a calculus of motives pushing a spiraling cognitive arms race into an intensely intentional world. The complexity of the social world that resulted may have become a selecting environment that put further pressures on cognitive development and interactional skills.

In sociocultural anthropology, as in the social sciences generally, the *interactional* foundations of social life have not been a central focus of research. Indeed, the key method of participant-observation presupposes the transparency of the interactional medium through which research is done, like the entomologist who uses the microscope as a tool rather than analyzing it. We trade on our “common humanity” to do anthropological research, yet typically without documenting or analyzing the mediating interactional interface. But if we ask how interaction itself works as a system, and how through the specifics of social interaction we can come to learn the things we know—both as analysts and as participants—important empirical questions arise. To what extent is the unexamined interactional system a constant across sociocultural space and time? To what extent is cultural context constructed by modulating specific interactional parameters? To what extent can differences in the conduct of social interaction affect cognitive or cultural categories?

Despite a relative neglect of the details of face-to-face interaction in mainstream sociocultural anthropology, our theme has numerous points of contact with several vital strands of ethnographic research. Work examining cultural conceptions of the person as agent and actor suggests that in some traditional societies there is an ideological reluctance to attribute thoughts and intentions to others (Shore 1998). (However, work on divination and religion proposes that these practices have bases in our interactional, intention-attributing instincts: Boyer 1994, 2002; Goody 1995; Zeitlyn 1995.) Different ideologies of intersubjectivity may be related to child-rearing practices, as explored in literature on language socialization (see Gaskins; Schieffelin and Ochs

1986). A succession of frameworks has arisen from within linguistic anthropology for thinking about constraints on verbal interaction in (culture-)specific settings (Duranti 2001; Gumperz and Hymes 1986; Hymes 1964), and about the invocation of such frames on the fly through “contextualization cues” (Gumperz 1982). This wider cultural context is captured in Geertz’s notion of “thick description” (Geertz 1973), tying the ideological and historical particularities of culture to the conduct of everyday life. Not unrelated in spirit is an important cross-fertilization between ethnographic research and the microanalysis of interaction (e.g., Goffman 1963, 1964, 1974; Goodwin 1994, 2000; Sidnell 2001, 2005), which connects through to the new “cognitive ethnography” traditions that examine distributed cognition, the idea that social institutions work through a cognitive division of labor organized in situated interpersonal interaction (Hutchins 1995).

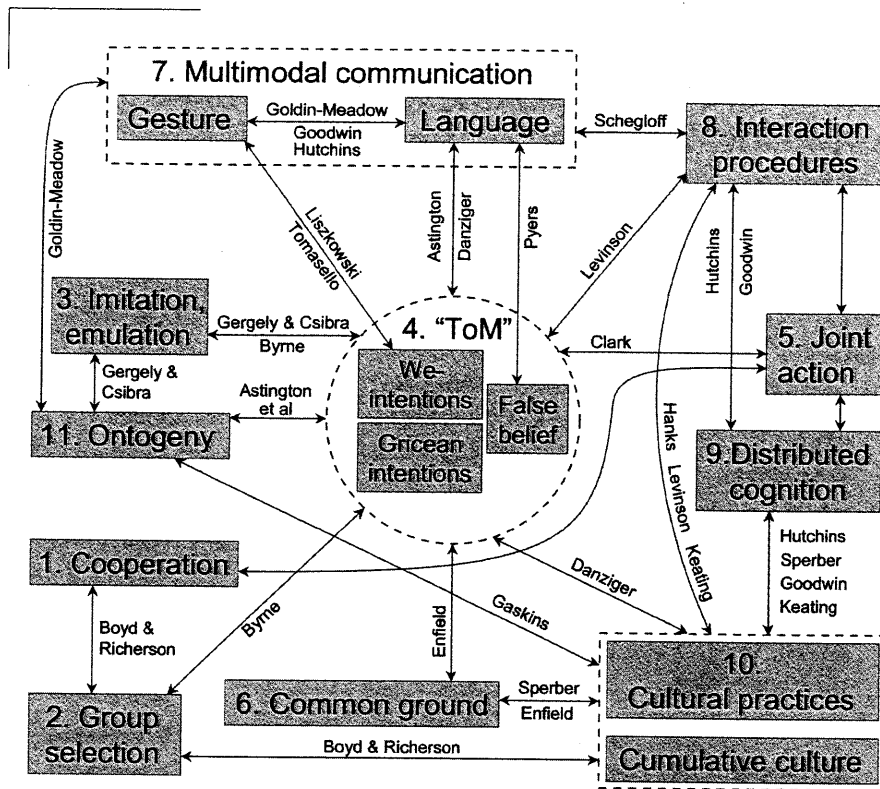
Notwithstanding these glimpses of insight from within social and linguistic anthropology—many of which are drawn on in the chapters of this book—the foundational nature of social interaction, in all its detail, is yet to be properly recognized in the larger compass of social anthropology.

### **Human Interaction as the Focus of a New Interdisciplinary Field**

One aim of this book is to define and consolidate a new field of research, a multidisciplinary approach to human interaction, its organization, and its constitutive role in social life. The project asserts the centrality of social interaction in the organization of human societies. Research in multiple disciplines shows just how intricately organized human interaction is, using multimodal channels of communication, building on detailed presumptions and shared understandings, foreseeing courses of action, and attuning to cultural settings. Underlying all this is a specialized cognition, crucially involving intention attribution or “mind reading” and the accumulation of shared understandings that makes historical culture possible.

From the chapters of this book, there emerges a set of closely interlocked concepts and lines of enquiry, which we sketch in Fig. I.1.

In the diagram, the boxes represent some of the crucial concepts in the discussion. The arrows show the links emphasized by different authors in this book. The diagram can stand as a mnemonic for the complex arguments adduced throughout the book, demonstrating interconnections between what at first seem rather different areas



**Figure 1.1.** Interlocking concepts developed in the different chapters. ToM = Theory of Mind.

of research. The density of connections indexes a clear domain of integrated inquiry.

To see this, let us take a single question, and follow where it leads us through the network of ideas and the chapters that make up this book. How did our unique brand of sociality evolve, with its varied linguistic and cultural environments? If we start with cooperation (box 1), we land in the thick of it. As discussed in the Boyd and Richerson chapter, the origin of cooperation is a deep puzzle in evolutionary theory. If this was social, a sort of enforced amity or Hobbesian contract, the answer might lie outside a theory of biological evolution. Yet the chapters by Gergely and Csibra, Tomasello, and Liszkowski each report very early cooperative acts by infants, suggesting the existence of cooperative instincts. Boyd and Richerson argue that this can only be accounted for by group selection (box 2). But this would rest on abilities to emulate

and imitate, the learning basis for building groupwide behavioral patterns (box 3). Byrne shows, observing gorillas, that these cognitive preconditions for cultural learning can have simple roots in behavior parsing. Gergely shows that human infants further analyze actions for their means–ends rationality, and will imitate “irrational” arbitrary (and thus potentially cultural) actions only when rational analysis fails. This presupposes the central ability to attribute intentions, or “mind read” (Theory of Mind, box 4). As a number of the chapters show, elements of ToM appear early enough in human development to suggest an instinctual basis. ToM abilities underlie joint action (box 5, see Clark), resting in part on cooperative instincts. Joint action must have been a crucial factor in the increased fitness of the group, which would have incrementally established biological foundations for ToM. Once our ancestors had developed some ToM, they would have had a basis for designing communicative actions in just such a way as to get others to recognize their intentions (Gricean intentions, box 4). This would then provide foundations for a rich kind of communication system without parallel in the animal world (see Levinson). These Gricean intentions rely for their recognition on keeping track of shared experience and knowledge (box 6), as discussed by Enfield. The exploitation of common ground depends on public signaling and display, which can exploit all the expressive modalities (box 7; see Goldin-Meadow), and indeed the props provided by the environment (see Goodwin, Hutchins). Even simple output systems can have rich interpretations, as demonstrated by Goodwin. This predicts that languagelike communication can arise without the provision of a conventional language, and this is what indeed happens in “home-sign” systems (see Goldin-Meadow). Once our ancestors had evolved a languagelike system of any advanced complexity, it would by a feedback relation have greatly amplified the power of ToM (see Astington, Pyers).

So far, we have considered a set of interlocked properties that the human individual brings to the task of conducting social interaction, which plausibly evolved together. But these properties are deployed in a highly structured system of interaction (box 8), with its rapid turn transitions, repair systems, overall sequential structure, and so forth (see Schegloff). These seem to have a universal, cross-cultural base, and are thus equally characteristic of human sociality. In fact, there are multiple connections between the interaction system and ToM. For example, the rapid turn-taking system and the repair system are the major guarantors of shared subjective understandings (Schegloff 1992). When what I say reveals a misunderstanding of what you

said, you get an immediate opportunity to correct it. An interaction system based on shared understandings provides an environment for distributed cognition (box 9), that is, for the distribution of cognitive labor that underlies effective joint action, as described by Hutchins. Such a system has emergent properties, shifting the burden of explanation away from properties of the individual to the shared activity. We can “read” other minds in part because human interaction is organized so as to engender intersubjective understanding. This fits the evolutionary theory sketched by Boyd and Richerson: We can only have evolved cooperative instincts in an environment in which joint action could endow a group with selective advantages.

Those selective advantages rest on the rapid adaptability of groups to circumstance. This entails cultural diversity (box 10). Social interaction has a distinctly different flavor across cultures, with the central media of communication—human languages—showing striking variation. Cultural learning in ontogeny (box 11), with its foundations in imitation (box 3), allows the accumulation of specific cultural practice, building the common ground that makes shared subjectivity possible. Cultural diversity in ideology and practice may feed back into local specializations of ToM and interaction practice (see Danziger, Gaskins, Hanks, and Pyers, in particular). Cultural systems are subject to their own evolutionary mechanisms (see Sperber) and to rapid change and adaptation (see Keating), allowing social groups to respond rapidly to new opportunities or challenges. This brings us full circle, back to the group selection that could have favored cooperative instincts in the first place.

What this exercise shows is that we will not obtain a good grasp of the evolutionary background to our species, and the unique properties of our social life, without understanding the links between these diverse aspects of our psychological and behavioral makeup. They form a web of interconnected properties that together constitute human sociality.

### **The Logic of this Book and its Organization**

The chapters of this book explore human sociality from a range of disciplinary perspectives. As the previous section will have made clear, this means that organizing them into a linear order is a challenge. Our aim in this section is to outline our solution (one of many possible), and to weave the threads into a structured whole.

### *Part 1: Properties of Human Interaction*

The most directly accessible manifestation of human sociality is face-to-face interaction, unfolding in real time, in conversation or some other type of sustained copresent engagement. What are its properties? The chapters in part 1 deal with the organization of interaction, touching on issues of copresence and engagement, sequence and intersubjectivity, and coordination and commitment.

*An Interaction Engine (Levinson)* Levinson begins with a bird’s-eye view of human behavior, arguing for a universal base to the species-specific way in which humans interact with one another. The underlying principles governing human interaction appear to be independent of specific languages or specific cultures. Indeed, they continue to operate where there is no shared language and culture. The language independence of these interaction principles, along with their facilitatory effect on language, suggests a phylogenetic priority of interaction principles over language in the history of the species.

What are these interaction principles? Levinson suggests that one can think of humans as being endowed with an *interaction engine*, consisting of a raft of motivations, cooperative tendencies, multimodal communication systems, and psychological endowments. A crucial ingredient is the mental equipment for Gricean communication—that is, the ability to recognize intentions based on signals whose formulation has been designed such that just those intentions be recognized. This motivates many of the properties of interaction, including the turn-taking machinery of verbal interaction, which effectively requires understandings to be immediately tested and displayed. Despite this universal base, interaction patterns can vary dramatically across cultures, as every traveler knows. Indeed, they must do so, because they are the carriers of culture. So how then does one reconcile these differences with a rich universal basis to interactional behavior? Levinson explores cross-cultural differences in the naming of persons, for example, under taboo restrictions, and shows how local cultural constraints can interact with universal principles (Enfield and Stivers, in press). The suggestion is that much cultural variation can be accounted for in terms of tweaking the interaction engine and the generic principles governing social interaction. The overall idea then is not that the interaction engine produces cross-cultural uniformity but that it provides generic constructional principles on which cultural diversity may be built, in human interaction.

*Generic Problems and Their Generic Solutions (Schegloff)* Schegloff's chapter picks up on this theme, focusing on the strikingly flexible yet precise organization of the sequential structures of human interaction. Schegloff reviews a set of candidate generic solutions to the basic problems of coordination and intersubjectivity in social interaction. For example, everywhere in the world, as far as we know, informal talk in conversation is organized using a precise and rapid turn-taking system; it is subject to repair or correction in similar ways; it exhibits paired utterances like questions and answers; it has recognizable openings and closings. Rapid alternation of turns at talking allows misunderstandings to become clear, and the turn-taking system is so organized as to allow them to be dealt with as near as possible to where they occur. A conversational repair system acts as the main guarantor of intersubjective understanding (Schegloff 1992; Schegloff et al. 1977), playing a crucial role in any kind of human interaction, regardless of the semiotic system employed. Generic interaction mechanisms of the kind reviewed by Schegloff are what make interaction without conventional language possible, as in the home-sign systems described in Goldin-Meadow's and Levinson's chapters, or the interaction with an aphasic man described in Goodwin's chapter.

*The Local Richness of Social Interaction (Goodwin)* Goodwin draws attention to the semiotically rich environment of human sociality, the intensive mutual copresence definitive of social interaction as we know it. Interactants provide and access information simultaneously from a great array of sources, including lexical items, grammatical constructions, prosody, deployment of gaze, facial expression, bodily comportment, and hand gesture. Even the "imperfections" of natural speech such as "errors" and their repair carry important information, and may be strategically managed. For example, a speaker may break off his or her own speech before completion to secure the visual attention (i.e., eye gaze) of another. The types of linguistic break-offs and restarts that result (and that pepper normal speech) have the effect of making explicit their internal syntactic structure, potentially providing an account for syntactic learning (cf. the discussion of action parsing in Byrne's chapter).

By Goodwin's account, an individual's production of talk and action is an intrinsically public, collaborative process. His case study of interaction with Chil, a severely aphasic man, draws into stark relief the kind of collaborative meaning making that is going on all the time in "normal" conversation. Chil's communication problems are overcome

in collaboration with his interlocutors, via common exploitation of semiotic resources in the immediate environment. (These "resources" include the other people in the interaction.) Despite having only three words and some use of gesture, Chil is able to engage successfully in social interaction. Again, as with home-sign systems used by deaf children (see Goldin-Meadow, Levinson), the human interactive system affords collaborative construction of meaning with very slender resources.

*Collaboration and Commitment (Clark)* Clark argues that the sort of focused, sustained interaction described by Levinson, Schegloff, and Goodwin presupposes individuals' commitment to the interaction as a collaborative activity. This is the kind of social commitment that makes it difficult to get off the telephone without first getting into (and through!) a closing sequence. Any kind of joint action requires mutual commitment, and has to be coordinated in some way. Consider the simple coordination of action involved in moving a table together. You have to pick up one end, and I the other, at more or less the same time, then I must move a bit, and you too, relating my speed to your speed. We have to know where we are going, and mutually monitor potential hazards like steps, and so on. Clark proposes that the commitments required to succeed in joint action are hierarchically organized, and that minor commitments (like "let's lift the table now") are subordinate to higher-level ones (like "let's get the table into the living room"). As his review of the extraordinary Milgram Experiments demonstrates, making a higher-level social commitment entails lower-level commitments we might not have foreseen. This is because committing to the larger activity means committing to its subcomponents. In turn, refusing to commit to subcomponents can mean renegeing on one's existing commitment to the entire activity. This leads us into the powerful emotional and moral dimensions of social life.

## *Part 2: Psychological Foundations*

The organization of social interaction and the commitments it entails presuppose special psychological underpinnings. The ability to recognize others' states of mind, whether attentional or volitional, and to share these states of mind through mutual focus in the ongoing course of interaction, is indispensable for human sociality. Chapters in the second part focus on the nature and development of the psychological basis for social interaction.



*Pointing, 1—The Ontogenetic Kernel of Human Sociality (Liszkowski)* Liszkowski reports on a series of experiments designed to explore the bases of infant pointing (see also Tomasello). As in the chapters by Goodwin, Levinson, and Goldin-Meadow, this gives us a sense of how social interaction can work without (full) language capacities. Prelinguistic infants make pointing gestures, but it has been unclear whether they are doing languagelike communication with this, as opposed to, say, spontaneously expressing their internal response to an object or event, or simply trying to get attention. Liszkowski gives evidence that 12-month-old infants point to communicate, taking into account other's goals and apparent knowledge states. If an experimenter appears to be looking for something he had a moment ago, but which has now gone out of his view, a one-year-old infant will point to it as a way of telling the adult where it is. Further, if the experimenter misunderstands an infant's pointing gesture, the infant will try again. This is a spectacular finding, because ToM literature standardly suggests that the ability crucial to this account (i.e., knowing that the other does not know something) is a much later achievement in development, coming not at 12 months but at four years. In Liszkowski's studies, the child is clearly using pointing for informing, one of the main motivations for communication. The children in these experiments are not only informing the adult experimenters but helping them. This is suggestive of early cooperative instincts (see Boyd and Richerson, Tomasello), particularly as the helping uses of pointing are employed here in interaction with people other than a main caretaker.

*ToM, 1—The Suite of Capacities and the Role of Language (Astington)* The phenomenon of pointing takes us more centrally to how "mind reading" may work. The foundations here are (1) having a grasp that others have mental states and (2) recognizing that these may diverge from one's own. This must involve an awareness of one's own kinds of mental states, and arguably the ability to employ such mental states in explaining the actions of others. Astington's chapter reviews what is known about children's development of such a ToM. A number of researchers (see Gergely and Csibra, Liszkowski, Tomasello) believe that human infants first grasp the nature of the other as an intentional agent from about nine months. But it is also widely accepted that a fully comprehensive ToM, as indicated by false belief understanding, is slow to mature, coming significantly after the full essentials of language are in place. Astington argues that language plays a key role. She reviews three ways in which this has been proposed in existing literature: knowledge and

use of mental state verbs with meanings like "want," "think," "know," and "believe"; knowledge and use of the complex syntactic structures associated with these mental state predicates; and firsthand experience of face-to-face conversation.

*ToM, 2—Consequences of Language Deficit (Pyers)* Pyers's chapter narrows in more closely on the relation between language and ToM, with a case study of a Nicaraguan sign-language community. Pyers's research reveals startling evidence for the crucial role that language may play in acquiring ToM capacities. In Nicaragua, a substantial Deaf population was only in the last decades brought together into a socially networked speech community, thanks to the establishment of educational institutions for the Deaf. This has led to the growth of a new natural language known as Nicaraguan Sign Language, a Creole born of many smaller home-sign or village-sign systems. The first generation of signers learned what was effectively a pidgin with limited expressive power. In addition, they were late learners of language in any form. By contrast, younger signers of the following generation have had the benefit of exposure to a developed sign language from a young age. Pyers reports that tests for ToM capacities show the younger signers to have a significant edge. The older signers do not master standard false belief tasks. This is *prima facie* evidence that language plays a determining role in the acquisition and application of ToM.

*Imitation and Rational Learning (Gergely and Csibra)* A developmental perspective on the question of how we read intentions into the actions of others is pursued by Gergely and Csibra. They investigate human infants' imitation of adults' actions, finding that infants do not just copy actions, but analyze the goal directness of others' behavior and look for the rationale behind the means chosen for carrying out an action, doing selective imitation accordingly. Thus, if a woman with her hands tied turns on a light with her head, an infant imitating this action will turn on the light with his or her hand (Gergely et al. 2002). This imitation achieves the same goal (getting the light to go on), but does not reproduce the means. The child surmises that the adult would have used her hands if she could have: that is, given that her hands were full, the woman's unusual action of using her head is rational. But in a different experimental condition, in which the adult has hands *free* and yet turns on the light using her head, the infant will use his or her head as well in imitating this. In this case, the woman *could have* used her hands to do the action, but does not. The child extracts a different

rationale for the marked manner of action, surmising that it was this unusual manner that was intended (i.e., here the adult *chooses* to use the head and not the hands), thus being a defining and not merely contingent part of the action the adult performed.

The possibility for rational learning of this kind is critical for the acquisition of culture. Cultural actions have both rational means–ends aspects (like collecting food and preparing it to eat) and nonrational, culturally constrained aspects (like eating with a knife and fork rather than with the fingers). Our children have to acquire both. In each case, the process of acquisition involves intention attribution based on direct observation of others' actions (cf. the description of action parsing in Byrne, and syntactic parsing in Goodwin). It happens that sometimes part of the goal of an action is that the action be done in a specific manner. This applies in the case of culturally stylized action.

### *Part 3: Culture and Sociality*

A number of chapters discussed so far touch on culture. In part 1, both Schegloff and Goodwin suggest that the nature and details of copresent engagement have a direct bearing on the establishment of a common worldview of interactants (see also Enfield). Levinson's chapter lays out the view that social interaction shapes, and is shaped by, local norms and routines. In addition, as Goodwin and Hutchins point out, the artifactual environment directly shapes our interactions. The cultural shaping of the material world can thereby feed back into the nature and organization of sociality. In part 2, Gergely and Csibra deal with one of the ways in which such cultural identification is signaled, providing the possibility for rational acquisition of locally specific (and "nonrational") manners of carrying out practical actions. The conventional manners people actually learn are locally defined, and historically emergent. Further, in part 5, as previewed below, the cooperative instincts of central importance to Boyd and Richerson are linked to a drive to maintain social identification with specific social groups, leading ultimately to the development of distinct cultures. Selective advantages of being cultural beings rest on the rapid adaptability of culture to circumstance, which underlies the cultural diversity that characterizes our world.

The four chapters in part 3 explore the relation of culture to human sociality, including implications of cultural variation.

*Local Ideologies of Intention Attribution (Danziger)* Danziger deals frontally with a question of cultural ideology in the analysis and interpretation

of human sociality. Given that notions such as intention attribution and ToM have been developed in Western society, to what extent do these notions reflect a bias in our own cultural practices? It is not known in what way or to what degree ToM has uniform cross-cultural relevance. There may be cultures with distinctly different ideas about the readability of others' intentions. Danziger explores a reluctance to attribute intentions beyond the literal content of what is said among the Mopan, a Mayan Indian group of Belize. The Mopan hold that a sincere statement that turns out to be false is a "lie," and they do not consider to be a "lie" an insincere statement that turns out by accident to be true. But they are quite able to pass false-belief tasks. It is not that the Mopan lack ToM. Rather, they place cultural limits on the inferences one may make from behavior, including speech. Danziger concludes that there are profound consequences of Mopan cultural ideology about the role of intention in meaning.

*Cultural Variation in Caregiver–Child Interaction (Gaskins)* Although most research on child development is conducted in Western settings, there are significant differences across cultures of the world in the ways in which caregivers and infants interact. These imply differences in the socialization of children into interaction itself. Gaskins points out a theoretical tension between two major presumptions in consideration of this issue. On the one hand, the essential outcomes of human cognitive development are presumed to be universal (see Astington, Tomasello). Children everywhere become competent adults. On the other hand, the processes that lead to successful socialization are dependent on very distinct kinds of input for learning in highly varied cultural settings. Gaskins's chapter offers a review and analysis of ethnographic research on child–caregiver interaction and socialization, detailing extreme variation in how infants are treated across cultures. The Western tendency for adult caregivers to try to induce focused interaction, with its attendant motherese, peek-a-boo routines and the like, appears to be exceptional. In many other cultures adult caregivers attempt to forestall needs and thereby preempt interaction. Sustained interaction with eye contact is rare. In Western cultures the emphasis is on the adult attempting to interpret the child's communications, whereas in many other cultures the onus is on the child to understand the adult. These striking contrasts raise questions about the universality of both the developmental process and its outcomes (see Pyers's evidence from Nicaraguan signers). Gaskins argues, however, that no matter how dramatic cultural variation appears to be, it must be providing an

environment in which certain fundamentals of sociality can develop. Some concrete possibilities are suggested in chapters exploring the early development of cognitive abilities critical to language and social intelligence (e.g., Gergely and Csibra, Liszkowski, Tomasello), including practices of establishing and maintaining joint attention (e.g., by finger pointing), sufficient to impart the ability for shared intentionality.

*Integrating Multiple Frames and Participant Roles (Hanks)* Socialization not only gives rise to general abilities and local retoolings of these but it brings a mountain of shared background for organizing and framing interaction in culture-specific ways. Culture supplies rich resources for participants to frame their engagements, and to adopt culturally relevant participant roles. In turn, interlocutors have to be able to recognize the specific frames and roles that are relevant to the interaction at hand. Further, as Hanks explores in his chapter, there may be multiple such frames and roles, as locally specified in a given cultural setting. This poses an integration problem for interactants. In addition, in many situations, interlocutors have to deal with distinct discrepancies in knowledge (or what Hanks dubs *uncommon ground*). Hanks explores these themes with reference to an extended example of shamanic curing sessions in Yucatan. These sessions have elaborate structure, with phases of conversational exchange between shaman and patient, phases of prayer, and phases in which the shaman addresses the divining crystals, discerning answers to his queries as if reading the minds of the spirits. There is a layering of interchange within interchange (talk to the spirits within talk to the patient), as well as a layering of cultural institutions. Hanks's chapter raises the challenge that all accounts of social interaction need to face: How do we integrate our general "social instincts" with specific cultural and often multilayered settings?

*Evolution of Cultural Convention Through Interaction (Keating)* Although cultural traditions like Mayan shamanism can be stable over millennia, they can also quickly evolve. Keating's chapter describes rapid adaptation of a conventional communication system (American Sign Language) to the new technology of videophone connections on computer. She shows that signers are fast establishing conventions from the new possibilities offered by the medium of communication. For example, one can move the hands forward for emphasis, placing them close to the camera such that they take up more of the visual field. (This would not work in real signing space, i.e., during face-to-face interaction.) Similarly, using the immediate feedback from the monitor of one's own signing as seen by

the interlocutor, one can exploit the collapse of the third dimension on the screen (e.g., pointing left to empty space so that it looks as if one is pointing to someone behind and to the left). Signing is hereby acquiring a new genre, with conventions of its own in the making. This case study shows rapid exploitation of new affordances offered by a change in the technological environment. Keating's observations dovetail with those of Goodwin and Hutchins concerning the key role of environmental affordances in human communication and cognition. It is this kind of potential speed of change in public conventions, compared with the glacial pace of genetic change, that gives both culture and the particular form of human sociality their adaptive value from an evolutionary point of view.

#### *Part 4: Cognition in Interaction*

Chapters in part 4 focus on cognition in interaction, and its consequences, cross-cutting the key concerns of parts 1–3—the organization of social interaction, its psychological underpinnings, and its sociocultural context(s). These chapters examine ways in which cognition and interaction not only interlock, but how they can be coconstitutive. The interactional setting is a primary context for the externalization of cognitive processes, where the relevant "cognitive artifacts" (Norman 1991) may include graphic devices, hand gestures, and the very people with whom we are interacting (see Goodwin). Such artifactual externalization of cognition can have both local and global effects, with consequences for the course of interaction itself, and for what ends up being shared among interactants as individuals in ongoing social relationships, and as common members of entire cultural systems.

*Making Thought Public, Without Language (Goldin-Meadow)* Goldin-Meadow explores how both symbols and thoughts emerge in interaction. First, she describes a striking example of communication working without conventional signs or symbols: the case of deaf children who are not exposed to a systematic conventional sign language, but instead construct a system of manual signs *de novo* (a so-called home-sign system; see Levinson). Sometimes, nondeaf parents of deaf children address them using spoken language only. In these cases, the child will invent a sign system of his or her own. The system is used one way, with the parent talking and gesturing back. These home-sign systems have languagelike properties: they show arbitrary form-meaning mappings; they are formally categorical; scenarios distant from the

here and now can be effectively described, and so forth. Such a system fundamentally relies on intention attribution (see Astington, Gergely and Csibra, Levinson), together with generic mechanisms for solving generic problems of communication and intersubjectivity (e.g., repair of nonunderstanding; see Goodwin, Schegloff). Goldin-Meadow's research shows how a species that had first evolved advanced interactional intelligence could, providing some cooperative instincts were in place, evolve a languagelike communication system. Here, many of the issues dealt with in the present book come together: the multimodality of social engagement, commitment and cooperation in social interaction, ToM and intention attribution, and emergence of convention.

The second section of Goldin-Meadow's chapter, focussing on gestures accompanying speech, shows how these freely inventive signals adumbrate "liminal" thoughts, allowing interactants to bring them into consciousness. Focusing on teacher-child interactions in arithmetic, in which children are still struggling to understand basic operations like subtraction, she finds that children unable yet to articulate or execute solutions, still betray a partial understanding in their gestures. Teachers unconsciously pick up on this inarticulate revelation of dawning comprehension, and can build their explanations on it. The hand betrays the thought, for gestures are cognitive artifacts (Enfield 2005b; see Hutchins), allowing communication in interaction to proceed where conventional language fails (as it did with the deaf children).

*Online Interaction and the Emergence of Structure (Hutchins)* Hutchins proposes another way in which the interactive system derives greater power than its structural components alone can contribute. His argument begins from a point emphasized in Goodwin's chapter, that there is a great deal of information publicly available in the environment of any given interaction. Environmentally coupled social interaction gives rise to a higher-level or emergent shared system of cognition. Social systems exploit this potentiality by structuring social activities such that they will have just these emergent effects. Hutchins argues that standard assumptions about the bases of social interaction, including ToM and intention attribution, overestimate what the individual brings to the task while underestimating what the task brings to the individual. This is amply illustrated in Hutchins's well-known example of what it takes to navigate a battleship into harbor (Hutchins 1995). As he explicates in his chapter, the navigation team on the bridge combine words and gestures with a map representing their path and position, so deciding on the bearings to use in navigating the massive vessel's course. The rest of

the calculations are automated as it were through the highly structured division of labor of the team and their instruments. Hutchins's point is that the entire overarching intelligence of the joint action cannot be attributed to any single individual. It is not represented in any single place but is emergent in the interactive activity. Hutchins suggests that the key to understanding human intelligence (including ToM) and its phylogeny (see part 5) is to see that higher-order cognition is first instantiated in joint activity. It thereby provides a selective environment for cognition about other minds, hence the development of ToM abilities.

*Building and Exploiting Common Ground (Enfield)* The possibility of rich interaction given scant semiotic resources, as described in chapters by Goldin-Meadow, Levinson, and Goodwin, is caused in great part by the presence of a massive inventory of common ground, both cultural and personal (Clark 1996). Common ground, or mutual knowledge shared by social associates (whether based on common experience or common cultural background) provides premises for amplicative inference (Goody 1995; Levinson 1995). Communication constantly exploits common ground, partly to overcome the communication bottleneck entailed by the slowness of speech (Levinson 2000). Enfield explores the notion that common ground is strategically exploited not only in the service of economy of expression, but for affiliative display of social closeness. Enfield suggests that because common ground is so crucial to effective communication, and to the display of affiliation, people go out of their way to augment it, as when a mother points out new things to her child yet without obvious or immediate purpose for doing so.

*From Individual Interactions via Cognition to Entire Cultural Systems (Sperber)* The chapters reviewed so far enable us to assemble a range of components of human sociality: its observable structures, its cognitive underpinnings, its cultural bases, and its role in the coordination of human cognition and activity. How are we to think about the link between cultural diversity and the presumably universal cognitive and ethological foundations of human sociality? A number of authors wrestle with this question, especially Gaskins, Levinson, Schegloff, and Astington. Sperber's chapter offers us a sustained theoretical panorama. He develops the idea of the Cognitive Causal Chain (CCC), a causal sequence that includes at least one cognitive representation. A perception is a causal relation between a thing in the world and a mental representation; an inference is a causal relation between two

representations; an action is a causal relation between an intention and the behavior that attempts to realize it. In social interaction the output of one individual's CCC is the input to another's, and in such cases we can talk about social CCCs. Great chains of social CCCs are possible, ultimately passing effects across whole populations. When these CCCs have the function of preserving either behavioral form or mental content or both (as in a song), they become cultural CCCs. This all leads toward a model of the distribution of cultural forms and meanings as if they were, say, viruses in a population—that is to say, subject to the mechanisms of evolution of traits in a population (Enfield 2003, 2005a; Sperber 1985, 1996). This suggests a Darwinian model for cultural evolution (Levinson and Jaisson 2006; Richerson and Boyd 2004). The relation between cognitive universals, provided by the organism, and the variability of cultural forms is simple enough: Cognition provides the essential filter on what can be easily transmitted through a CCC. To get feedback from CCCs to the cognitive system requires a further kind of evolutionary mechanism, which leads us to the chapters in the final part, focusing on the phylogeny of human sociality.

### *Part 5: Evolutionary Perspectives*

The chapters in parts 1–4 establish defining properties of human social interaction, including sustained coattentional engagement, common commitment to cooperative activity, and attribution of communicative intentions to others. These properties are not shared by even our closest relatives among the apes. What are the critical differences? How could they have evolved?

*Cooperative Instincts and Group Selection (Boyd and Richerson)* The mutual commitment characteristic of human interaction (see Clark, Goodwin) points to a classic puzzle in evolutionary theory: the riddle of human cooperative behavior. Why are people so highly cooperative, when, for an individual, it should always pay to take the benefits of others' cooperative acts without reciprocating? The answer supplied in Boyd and Richerson's chapter is that cooperative behavior is instinctual. (This is supported by work presented in a number of other chapters: Gergely and Csibra, Liszkowski, and Tomasello report cooperative acts by infants of around one year of age.) Boyd and Richerson discuss experimental findings that adults, from societies of different kinds around the world, do not maximize their own gains but, instead, feel an obligation to share hidden benefits (Henrich et al. 2004). If our brand of cooperation is a

species-specific instinct, we then face the evolutionary puzzles: What would have been the selective advantage of cooperative sociality for the individual? How did the mechanisms that drive it develop?

Boyd and Richerson argue that group selection provides an account for the evolution of human cooperative instincts. Group selection is an unusual mechanism for evolutionary change, in which behavior shared by a group, rather than by an individual or his or her immediate kin, gives the entire group advantages over other groups. Because of its marginal status as an evolutionary mechanism, group selection presupposes earlier cultural adaptations that would have given sufficient adaptive advantage to the group as a whole as well as behaviors that signal and maintain boundaries between groups. Thus, the cognitive prerequisites for cultural learning (see Byrne, Gergely and Csibra, Tomasello) would have been essential for the evolution of cooperative instincts.

*Evolution of Action Parsing and Intention Attribution (Byrne)* Many of the chapters in this book detail the structure and nature of interaction, showing that humans depend in interaction on the ability to segment and interpret complex and sustained sequences of action, to recognize routines within them, and to see the intentions behind them. How could such skills have developed in our species? Cultural learning clearly involves the ability to learn from watching others' behavior. But where this behavior is of any complexity, some kind of parsing analysis is required. Byrne describes how some groups of gorillas share techniques for nettle stripping that are transmitted by cultural learning. He proposes that a simple statistical and structural analysis of observed behavior allows a novice not only to extract the essentials of the technique (see the parallel account for syntactic parsing of speech in Goodwin's chapter), but to grasp its goal-oriented nature (cf. Gergely and Csibra's discussion of rational imitation). This provides a glimpse into the phylogenetic precursors of intention attribution, human imitation, and learning. It also has implications for our understanding of intention attribution in modern human social interaction.

Like a number of other contributors (see Goodwin, Hutchins, Schegloff), Byrne cautions against overestimating the degree to which people explicitly model others' mental states in interaction. The explicit mentalism implied by much ToM research can be minimized by behavior-based, statistical means for interpretation of others' action.

*Pointing, 2—The Phylogenetic Kernel of Human Sociality (Tomasello)* Several contributors discuss the importance of pointing in human interaction.

Goodwin, for example, describes the critical role of pointing in communicating when language is unavailable (see also Goodwin 2003). Tomasello's chapter puts the theme in a phylogenetic perspective. He starts from the observation that apes, our nearest relatives, not only lack language but they do not point or comprehend pointing. (This claim has been contested—De Waal 2001; Veà and Sabater-Pi 1998—but as Tomasello points out, the reported empirical observations have not been replicated; cf. Povinelli et al. 2003.) Although apes do monitor others' eye gaze, and seem to understand that others might see what they cannot see, they do not seem to grasp the idea that an interactant might be trying to get them to shift their attention. Underlying this failure is the absence of efforts to establish joint attention, and the absence of complex collaborative action. Experiments by Tomasello and colleagues show, by contrast, that human infants of 14 months systematically distinguish between what an adult has already seen from what is new for that adult. Tomasello argues that what distinguishes humans from other apes are instincts for helping and sharing, manifest in collaborative interactions based on "shared intentionality" (i.e., joint intentions and joint attention). These instincts are manifest in the humble pointing gesture, which despite being well under control by a one-year-old human (see Liszkowski), is never convincingly comprehended by any other great ape. It is this gesture, Tomasello submits, that provides a foundation for the evolution and acquisition of language, culture, and the full richness of human sociality.

### **Concluding Remarks: Toward a Synthesis**

#### *A Framework for Integrating the Different Levels of Phenomena in the Domain*

Here, we propose a synthesis of the ideas aired in this book, to show how the contributing concepts, which relate to distinct levels of phenomena, fit together to yield an integrated perspective on human sociality. The framework helps us see the essential roles different disciplines play in the study of this domain, and how they might better inform one another in future work. We distinguish three distinct levels of phenomena:

1. **Interaction Engine (individual level):** The individual brings to interaction an "interaction engine," consisting of ToM abilities and communicative capacities built on them, biological constraints, and ethological proclivities (as outlined in Levinson and in part 2 of this book). Crucial elements of this include the ability to recognize others' intentions through modeling the minds of others in real contexts

(and to anticipate *their* modeling of *our anticipation* of their intention attribution!). These elements together form the essential equipment for formulating and interpreting actions in an interactional setting. We think it likely that the foundations of the engine are biologically endowed, or at least unfold in human development in comparable ways given local parallels in interactional organization. But such development depends on an interaction matrix, so that the engine may be fine-tuned to a local cultural frame.

2. **Interaction Matrix (interpersonal level):** The "interaction matrix" in which the interaction engine is deployed has special and peculiar emergent properties, potentially accounting for the universality of its inventories of turn-taking systems, repair mechanisms, sequential organizations, and the like (issues explored largely in parts 1 & 4). An interaction is a sequential, contingent structure in which what happens next is as much determined by other parties as by oneself. There need not be any particular prearranged plan or direction (as in casual conversation). As yet, there is no adequate formal theory of this kind of contingent interaction with shifting goals (despite game theory being able to capture situations in which goals are zero-sum or fully shared). The most complex properties of human interaction are emergent. Consider a soccer team working together. The overall flow of movement of the ball stems from the individual players' movements and local intentions, but the entire pattern cannot be coherently reduced to any one player's individual intentions, tacit understandings, or actions. The emerging pattern depends on actual outcomes, overall sequence and timing.

3. **Sociocultural Frame (social-cultural level):** The interaction matrix provides the building blocks of social organization and its constituent institutions, which constrain interaction within specific cultural frames (focal in part 3). These are the frames in which the business of society is conducted, whether they are legal hearings, gossip on the street corner, or infant-caretaker settings. Social institutions are often robust, with deep histories, and their fortunes are subject to patterns of cultural evolution on a time scale different from the ephemeral interactions that nevertheless instantiate them. This is the level to which the bulk of ethnographic description and analysis has been devoted.

#### *Consequences of this Framework*

Within this framework, we can restate a number of interesting propositions arising out of the work summarized in this book. Consider the following points.

*On Human Phylogeny* A central key to human evolution lies in understanding the relation between phenomena of these three different scales and ontological types (central issues in part 5). The interaction engine is adapted to the interaction matrix, for the engine's function is to conduct mentally mediated interaction at the interpersonal level. The interaction matrix of our forebears was the selecting environment for the biological and ethological roots of the interaction engine—for example ToM, or the foundations of language. In turn, the interaction matrix is built out of the raw potential that the engine supplies. Limits to, say, speed or complexity of communication are inherited from the engine's properties. Again, the interaction matrix is adapted to conducting the business of higher-order social organizations. Among the properties of these higher-order units are those that endowed groups with the adaptive cultural edge over other groups, allowing group selection to play a role in human evolution.

*On Language* Language plays a central role in human social life, as suggested by its ubiquity, dominance, and elaboration. But the work assembled in this book suggests that language itself rests on other abilities that are ontogenetically, phylogenetically, and logically prior—in particular, the ability to attribute action, meaning, and intention in structured sequences of interaction. Communication is possible without fully fledged language (as in home-sign systems or in interaction with infants), operating on a basis of reflexive or anticipatory intention attribution, which is always at work even in the use of languages with full expressive power. Thus, the evolutionary basis for language must be sought in the mutual adaptation of the interaction engine to the interaction matrix, and to the sociocultural level in which it is embedded. To be sure, the combination of the interactional engine and a full preconventionalized symbolic system like a language yields a quantum leap in expressive and computational power in the interactional domain. The structured representational system of language also appears to retool the ToM (see Astington, Pyers), allowing richer and more complex inferences. But in the end, it seems, although language is transformative of our cognitive and interactional powers, it rests on a more fundamental cognitive specialization that appears earlier in human phylogeny (and ontogeny; see below).

*On Cultural Evolution* Cultural diversity arises from the relative success of particular institutions in local contexts, together with random effects like drift (see Sperber). Particular cultural patterning of interaction reflects feedback to the interaction matrix from specific forms of organization

in the sociocultural frame, and then to the local tuning in ontogeny of the interaction engines of individuals. For example, in a society like Java with social hierarchy and courtly traditions, decorum will specify the proper deployment of body position, gesture, and honorific levels in language, behavior inculcated during child development.

*On Human Ontogeny* The phylogenetic and historical perspectives are complemented by an ontogenetic perspective. The interaction engine is not literally delivered with the infant at birth, although core biological and ethological constituents certainly are. The engine has to unfold through experience in the interaction matrix, which will cause the developing child's interaction engine to inherit cultural specializations. Gaskins's catalogue of cultural differences in child rearing suggests, however, that the initial ingredients are robust enough to give us universal outcomes regardless of experience (as in the cross-cultural parallels in Goldin-Meadow's home-signing children). We suggest that culture can reach deep down into the details of interaction, but only by modulating tendencies that are universal or default.

## Conclusion

The kind of synthesis we propose offers a closer integration of the contributing research traditions. So psychological approaches will benefit from expertise at the level of the interaction matrix. For example, work on infant pointing gestures (see Liszkowski, Tomasello) should be alert to the sequential contexts in which they occur and on which their interpretation may crucially depend. Conversely, work on the interaction matrix will be enriched by understanding what is (psychologically) under the hood. Observational work on sequences of interaction has revealed many kinds of contingencies between actions in interaction (e.g., question–answer sequences or greetings), but we do not know how some of these implicit classifications (e.g., of an action as an X or a Y) are achieved online. We know little about the sources and development in infancy of skills in navigating finely temporal and contingent interactional sequences such as conversational turn taking. Do such skills have an instinctual basis, or are they built during development on a more primitive instinctual testing of contingencies in the physical world? We know that interactants are highly sensitive to others' mental states, but we do not know how these registers of information for potential interlocutors are constructed or assessed—experimental techniques will be critical here.

At another level, that of the sociocultural frame, the interaction matrix offers insights into how cultural events and processes are actually constructed. Slight modifications of a universal generic base for conversational organization can yield all sorts of specific speech events. For example, restricting interchanges to questions and answers can give us a basis for courtroom interrogation or classroom teaching—further assigning rights to question, and the role of overhearers, can help us distinguish the conduct of the two cultural event types. Tracing further back, if we know the psychological or developmental sources of those universal tendencies, we might understand universal constraints on social organization. Conversely, the analysis of social organization can inform the conduct of interaction in myriad ways, helping us understand background assumptions operative within specific events, the choice of language and social role, and the like.

This raises an apparent tension in this volume between those who emphasize the individual's psychological abilities and those who focus on the emergent properties of the interaction matrix, or the way in which social interaction is adapted to local sociocultural organization. We do not regard this as simply border warfare, with rival definitions of Durkheim's "psychological" versus "social facts." Rather, it reflects a disagreement about the primacy of one or other of the three levels—the individual, the interactional, and the sociocultural. When A asks B a question, and B answers it, is this because B discerns A's intentions (a psychological level of explanation)? Is it because B follows the rules of the language game (an interactional level of explanation)? Or is it because B recognizes that A is endowed with the social rights and authority to ask that kind of question in the current situation (a sociocultural level of explanation)? Different researchers rightly test the power of their own lines of explanation by pushing the limits, and they are likely to favor one or another level of explanation. This area of research is young enough that there is no consensus about which level should bear the major burden of explanation for specific phenomena. Thus, although there are substantive concerns raised in some of the chapters regarding the applicability of terms and concepts like "intention," "action," and even "cognition," true reduction to just one level or another is not going to work: the levels have independent properties but are also mutually interdependent. The interpersonally emergent interaction matrix would not be possible without the individually seated interaction engine, but it is not "generated" by it. The interaction matrix has higher-order emergent properties, reflected in the way that local outcomes are contingent on the actions and responses of all the players. Likewise,

although social institutions are realized through interaction, they have long-term historical roots and interdependence with other aspects of culture that require an independent level of analysis. For these reasons, this will remain an interdisciplinary domain of inquiry, requiring input from disciplines with insights special to the different levels that make it up. And the contributors to this project will need to learn each others' languages if we are going to make real progress.

We thus bring to a close our preview of the range of ideas on human sociality put forth in the chapters of this book. We hope the volume does much to spur cross-border commerce between the different fields. If this can be promoted, we believe that the field of social interaction research will rightly come to be central in the human sciences, opening fundamental insights into what kind of a beast we are, and how we came to have our own uniquely complex form of sociality.

## Notes

1. The term *sociality* is used with a narrower meaning than ours by Henrich et al. (2004), to refer to cooperative and altruistic instincts, which "deviate from an axiom of selfishness." Sussman and Chapman (2004) use the term in a related way to this, to refer to the orientation of individuals to group living. Given that "group-living individuals must forgo some of their individual freedoms in order to socialize within the 'group,'" Sussman and Chapman's sense of "sociality" refers to "the compromises that individuals make, the mechanisms they use, and the means by which they maintain these social groups" (Sussman and Chapman 2004:10). Our sense of sociality includes these features among a broader complex of psychological and social predispositions, principles of interactional organization, and specific interactional practices.
2. Key references include Premack and Woodruff (1978), Byrne and Whiten (1988), Astington et al. (1988), Davies and Stone (1995a, 1995b), Whiten and Byrne (1997), and Carruthers and Smith (1996), among many others.
3. Our use of the term *Theory of Mind* refers more generally to the full ensemble of "mind-reading" skills of which false-belief understanding is a single and late-developing component.
4. Key references include Sacks (1992), Sudnow (1972), Sacks et al. (1974), Goodwin (1981), Atkinson and Heritage (1984), Button and Lee (1987), Schegloff (in press), among many others.



5. Interaction analysts have also invested effort in understanding the use of gesture, gaze, and body position in social interaction (Goodwin 1981; Schegloff 1984; see also Goodwin, Hutchins). (Psychologists, too, have been especially interested in gesture; see Goldin-Meadow, Liskowski, Tomasello.) These studies underline the multimodal nature of human communication. Again, there are clear universal tendencies here. For example, in all cultures, as far as we know, people gesture when they talk, although the exact nature of gesture, gaze, and body position are very much culturally constrained.

## References

- Astington, J. W., P. L. Harris, and D. R. Olson (eds.). 1988. *Developing Theories of Mind*. Cambridge: Cambridge University Press.
- Atkinson, J. M., and J. Heritage (eds.). 1984. *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.
- Baron-Cohen, S. 1995. *Mindblindness: An essay on autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Boyd, R., and P. J. Richerson. 2005. *The origin and evolution of cultures*. New York: Oxford University Press.
- Boyer, P. 1994. *The naturalness of religious ideas: A cognitive theory of religion*. Berkeley: University of California Press.
- . 2002. *Religion explained: The human instincts that fashion gods, spirits, and ancestors*. London: Vintage.
- Button, G., and J. R. E. Lee (eds.). 1987. *Talk and social organization*. Clevedon, UK: Multilingual Matters.
- Byrne, R. W., and A. Whiten (eds.). 1988. *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Clarendon Press.
- Carruthers, P., and P. K. Smith (eds.). 1996. *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Clark, H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Davies, M., and T. Stone (eds.). 1995a. *Folk psychology*. Oxford: Blackwell.
- , (eds.). 1995b. *Mental simulation*. Oxford: Blackwell.
- De Waal, F. 2001. Pointing primates: Sharing knowledge without language. *Chronicle of Higher Education*, January 19: B7–B9.
- Dunbar, R., C. Knight, and C. Power (eds.). 1999. *The evolution of culture*. New Brunswick, NJ: Rutgers University Press.
- Duranti, A. (ed.). 2001. *Linguistic anthropology: A reader*. Malden, MA: Blackwell.

- Enfield, N. J. 2000. The theory of cultural logic: How individuals combine social intelligence with semiotics to create and maintain cultural meaning. *Cultural Dynamics* 12(1):35–64.
- . 2003. *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia*. London: Routledge.
- . 2005a. Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology* 34:181–206.
- . 2005b. The body as a cognitive artifact in kinship representations: Hand gesture diagrams by speakers of Lao. *Current Anthropology* 41(6):51–81.
- Enfield, N. J., and Stivers, T. (eds.). in press. *Person reference in interaction: Linguistic, cultural, and social perspectives*. Cambridge: Cambridge University Press.
- Geertz, C. 1973. *The interpretation of cultures*. New York: Basic Books.
- Gergely, G., Bekkering, H., and Király, I. 2002. Rational imitation in preverbal infants. *Nature*, 415(6873):755.
- Goffman, E. 1963. *Behaviour in public places: Notes on the social organization of gatherings*. New York: Free Press.
- . 1964. The neglected situation. *American Anthropologist* 66(6):133–36.
- . 1974. *Frame analysis: An essay on the organization of experience*. Boston: Northeastern University Press.
- . 1981. *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, C. 1981. *Interactional organization: Interaction between speakers and hearers*. New York: Academic Press.
- . 1994. Professional vision. *American Anthropologist* 96(3):606–633.
- . 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32:1489–1522.
- . 2003. Pointing as situated practice. In *Pointing: Where language, culture, and cognition meet*, edited by S. Kita, 217–242. Mahwah, NJ: Erlbaum.
- Goody, E. N. (ed.). 1995. *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence*. Cambridge: Cambridge University Press.
- Grice, H. P. 1957. Meaning. *Philosophical Review* 67:377–388.
- . 1975. Logic and conversation. In *Speech Acts*, edited by P. Cole and J. L. Morgan, 41–58. New York: Academic Press.
- Gumperz, J. J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

- Gumperz, J. J., and D. Hymes (eds.). 1986[1972]. *Directions in sociolinguistics: The ethnography of communication*. London: Blackwell.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, and H. Gintis (eds.). 2004. *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. Oxford: Oxford University Press.
- Hutchins, E. 1995. *Cognition in the wild*. Cambridge, MA: MIT Press.
- Hymes, D. H. (ed.). 1964. *Language in culture and society: A reader in linguistics and anthropology*. New York: Harper and Row.
- Kockelman, P. 2005. The semiotic stance. *Semiotica* 157(1-4):233-304.
- Levinson, S. C. 1995. Interactional biases in human thinking. In *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence*, edited by E. Goody, 221-260. Cambridge: Cambridge University Press.
- . 2000. *Presumptive meanings*. Cambridge, MA: MIT Press.
- Levinson, S. C., and P. Jaisson (eds.). 2006. *Evolution and culture*. Cambridge, MA: MIT Press.
- Norman, D. A. 1991. Cognitive Artifacts. In *Designing interaction: Psychology at the human-computer interface*, edited by J. M. Carroll, 17-38. Cambridge: Cambridge University Press.
- Povinelli, D. J., J. M. Bering, and S. Giambone. 2003. Chimpanzees' "pointing": Another error of the argument by analogy? In *Pointing: Where language, culture, and cognition meet*, edited by S. Kita, 35-68. Mahwah, NJ: Erlbaum.
- Premack, D., and G. Woodruff. 1978. Does the chimpanzee have a Theory of Mind? *Behavioral and Brain Sciences* 1:515-526.
- Richerson, P. J., and R. Boyd. 2004. *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Sacks, H. 1992. *Lectures on Conversation*. London: Blackwell.
- Sacks, H., and E. A. Schegloff. 1979. Two preferences in the organization of reference to persons in conversation and their interaction. In *Everyday language: Studies in ethnomethodology*, edited by G. Psathas, 15-21. New York: Irvington.
- Sacks, H., E. A. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4):696-735.
- Schegloff, E. A. 1984. On some gestures' relation to talk. In *Structures of social action: Studies in conversation analysis*, edited by J. M. Atkinson and J. Heritage, 266-296. Cambridge: Cambridge University Press.
- . 1992. Repair after next turn: The last structurally provided defense

- of intersubjectivity in conversation. *American Journal of Sociology* 97(5):1295-1345.
- . in press. *Sequence organization in interaction: A primer in conversation analysis, 1*. Cambridge: Cambridge University Press.
- Schegloff, E. A., G. Jefferson, and H. Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53(2):361-382.
- Schieffelin, B. B., and E. Ochs (eds.). 1986. *Language socialization across cultures*. Cambridge: Cambridge University Press.
- Shore, B. 1998. *Culture in mind: Cognition, culture, and the problem of meaning*. Oxford: Oxford University Press.
- Sidnell, J. 2001. Conversational turn-taking in a Caribbean English Creole. *Journal of Pragmatics* 33(8):1263-1290.
- . 2005. *Talk and practical epistemology: The social life of knowledge in a Caribbean community*. Amsterdam: Benjamins.
- Sperber, D. 1985. Anthropology and Psychology—Towards an epidemiology of representations. *Man* (n.s.) 20(1):73-89.
- . 1996. *Explaining culture. A naturalistic approach*. Oxford: Blackwell.
- Sperber, D., and D. Wilson. 1986. *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.
- Sudnow, D. (ed.). 1972. *Studies in social interaction*. New York: Free Press.
- Sussman, R. W., and A. R. Chapman. 2004. The nature and evolution of sociality: Introduction. In *The origins and nature of sociality*, edited by R. W. Sussman and A. R. Chapman, 3-22. New York: Aldine de Gruyter.
- Tomasello, M., M. Carpenter, J. Call, T. Behne, and H. Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28:675-735.
- Veà, J., and J. Sabater-Pi. 1998. Spontaneous pointing behaviour in the wild Pygmy Chimpanzee (*Pan paniscus*). *Folia Primatologica* 69:289-290.
- Whiten, A., and R. W. Byrne (eds.). 1997. *Machiavellian intelligence II: Extensions and evaluations*. Cambridge: Cambridge University Press.
- Zeitlyn, D. 1995. Divination as dialogue: Negotiation of meaning with random responses. In *Social intelligence and interaction: Expressions and implications of the social bias in human intelligence*, edited by E. N. Goody, 189-205. Cambridge: Cambridge University Press.

WENNER-GREN INTERNATIONAL SYMPOSIUM SERIES

Series Editor: Leslie C. Aiello, President, Wenner-Gren Foundation for Anthropological Research, New York.

ISSN: 1475-536X

*Previous titles in this series:*

*Anthropology Beyond Culture*

Edited by Richard G. Fox & Barbara J. King, 2002

*Property in Question: Value Transformation in the Global Economy*

Edited by Katherine Verdery & Caroline Humphrey, 2004

*Hearing Cultures: Essays on Sound, Listening and Modernity*

Edited by Veit Erlmann, 2004

*Embedding Ethics*

Edited by Lynn Meskell & Peter Pels, 2005

*World Anthropologies: Disciplinary Transformations within Systems of Power*

Edited by Gustavo Lins Ribeiro and Arturo Escobar, 2006

*Sensible Objects: Colonialisms, Museums and Material Culture*

Edited by Elizabeth Edwards, Chris Gosden and Ruth B. Phillips, 2006

Since its inception in 1941, the Wenner-Gren Foundation has convened more than 125 international symposia on pressing issues in anthropology. These symposia affirm the worth of anthropology and its capacity to address the nature of humankind from a wide variety of perspectives. Each symposium brings together participants from around the world, representing different theoretical disciplines and traditions, for a week-long engagement on a specific issue. The Wenner-Gren International Symposium Series was initiated in 2000 to ensure the publication and distribution of the results of the foundation's International Symposium Program.

Prior to this series, some landmark Wenner-Gren volumes include: *Man's Role in Changing the Face of the Earth* (1956), ed. William L. Thomas; *Man the Hunter* (1968), eds Irv DeVore and Richard B. Lee; *Cloth and Human Experience* (1989), eds Jane Schneider and Annette Weiner; and *Tools, Language and Cognition in Human Evolution* (1993), eds Kathleen Gibson and Tim Ingold. Reports on recent symposia and further information can be found on the foundation's website at [www.wennergren.org](http://www.wennergren.org).



# Roots of Human Sociality

Culture, Cognition and Interaction

*Edited by*

N. J. ENFIELD AND STEPHEN C. LEVINSON

2006

 **BERG**

Oxford • New York