Over the past three decades the major obstacle for scientists interested in the psychological aspects of human linguistic competence has been generative grammar. Throughout this period, generative grammarians have claimed, and many psychologists have believed, that the only interesting aspect of language is its syntax, and that syntactic structure consists wholly of mathematical algorithms that are independent of meaning, communicative intention, and other psychological processes (e.g., Chomsky, 1986; Pinker, 1994). Recently, however, a new linguistic paradigm has emerged, and it is much more congenial to the traditional concerns of psychologists. The paradigm of cognitive linguistics, and its companion functional linguistics, is explicitly committed to describing and investigating linguistic competence in psychologically meaningful terms (Lakoff, 1990). Cognitive linguists thus describe linguistic structure in terms of such things as symbols, categories, schemas, perspectives, images, communicative functions, and a variety of other fundamentally cognitive and social processes (E. A. Bates & MacWhinney, 1989; Lakoff, 1987; Langacker, 1987, 1991; Talmy, 1988; van Valin, 1993). With the descriptions of linguistic structure emerging in this new paradigm, we have, for the first time, the
possibility of creating a real psychology of language (Tomasello, 1992a, 1992b).

Cognitive linguists have, for the most part, focused their research attention on the conceptualizations that particular types of linguistic expressions encode and how they encode them, for example, the way languages structure space, time, events, agency, and so forth (see, e.g., the papers in Rudzka-Ostyn, 1988). Functional linguists have focused mostly on very general communicative functions such as topic and comment and how these are encoded in various linguistic structures in the many languages of the world (e.g., the papers in Haiman, 1985, and van Valin, 1993). There is another dimension to the psychology of language that these researchers have mostly ignored, however, and that is the cultural context within which languages arise, both phylogenetically and ontogenetically. Over and above a concern with the kinds of things human beings talk about and how they talk about them, it is also important to study the processes of cultural cognition, interaction, and learning that make the acquisition and use of language possible in the first place.

My goal in this chapter is to spell out these cultural processes in some detail, with special reference to the ontology of language. Throughout, it should be kept in mind that there are two complementary aspects to each of these processes. On the one hand are the cognitive and social-cognitive processes that constitute the human adaptation for culture: Human children bring to language acquisition a number of social-cognitive and cultural learning skills by means of which they are able to comprehend and reproduce for themselves adult acts of linguistic communication. Especially important in this regard are children's developing skills in understanding the intentional actions of other persons, including their acts of linguistic reference. On the other hand are the cultural structures and institutions that exist prior to each child's birth: Human adults are adapted to pass on cultural skills to their children. Of special interest in this case are the routine cultural activities and other joint attentional interactions that adults structure for young children, and within which they treat children as intentional agents and use much of their early child-directed speech. Such cultural settings and interactions are an essential ingredient in children's coming to understand the intentional acts and cultural activities of other persons, including language.

In this chapter, I explore these two aspects of language acquisition at each of the two main levels of structure that distinguish language from other forms of animal communication: the level of individual lexical symbols and the level of grammatical symbols. In each case I provide some phylogenetic perspective on the process by making explicit comparisons to the communicative skills of our nearest primate relatives, the chimpanzees—both those that have developed in species-typical environments and those that have developed in human-like cultural environments. I conclude with some speculations on the evolution of culture and language.

LEXICAL SYMBOLS

Deacon (chap. 5, this volume) argued that, despite the great attention paid to grammar as the unique characteristic of human language, linguistic symbols are also uniquely human (at least when we consider nonhuman animals in their natural habitats), and indeed a strong argument can be made that symbols are the key to understanding all of human language. This is certainly the conclusion reached by researchers attempting to teach human-like linguistic skills to apes, as it has quickly become apparent that stringing together communicative behaviors is not grammatical if those behaviors are only conditioned responses and not true symbols (Savage-Rumbaugh, 1986). Even more radically, linguists such as Langacker (1987, 1991) have argued that the grammar of a language not only rests on symbols, but is, in reality, just another form of symbolic functioning: Grammatical symbols are simply symbols that may be used to designate certain kinds of relational or structural meanings. It is thus crucially important to begin any analysis of language by determining something of the nature of symbols and how they might arise both in phylogeny and ontogeny.

In recent thinking about linguistic symbols, two characteristics primarily have been emphasized. From the point of view of cognitive psychology, symbols serve an "information representation" function (Huttenlocher & Higgins, 1978; Premack, 1990). From the point of view of developmental psychology, symbols are communicative behaviors that have become "decontextualized" in that they can be used, not just as a part of an interactive sequence, but to "stand for" it (E. A. Bates, 1979;
Werner & Kaplan, 1963). These two functions of linguistic symbols are clearly an important part of the story. For current purposes, however, I would like to emphasize two related, but different, characteristics of symbols that bring out even more clearly their inherently social-cultural nature. First, of primary importance is the use of linguistic symbols in acts of external reference and, especially, acts of predication (Reed, 1993). In the analysis of Tomasello (1995), an act of linguistic reference is an act in which one individual intends that another individual should attend to some aspect of their shared external environment, in which an act of attention is regarded, according to E. J. Gibson and Rader (1979), as an act of intentional perception. This translates thus: A intends for B to intentionally perceive X. An act of predication simply extends this process as one individual first secures joint attention to some entity with another individual and then expresses an intention that the other should attend to one aspect, of many possible aspects, of that entity.

The second key characteristic of symbols is their reciprocal quality in being communicative behaviors that people both understand and understand that others understand (Mead, 1934). In the analysis of Savage-Rumbaugh (1990), a linguistic symbol is a communicative act that the producer comprehends—in some sense from the perspective of the receiver—as she produces it. This Janus-like quality is what gives linguistic symbols their special role in the evolution and transmission of culture, as both adult and child communicate with each other using the same system of symbols (Herford, 1989).

These two characteristics of linguistic symbols highlight clearly their social-cultural nature, as neither can be described without reference to at least two individuals engaged in certain kinds of social-cultural interactions. For purposes of exposition, these two characteristics of linguistic symbols are brought out even more clearly if we compare language to other forms of communication—more specifically, if we compare the communicative competencies of chimpanzees in their natural habitats, human children in the early phases of language acquisition, and chimpanzees raised in human-like cultural environments. I treat these each in turn, with some concern in each case for both the role of social-cultural cognition (especially the understanding of intentions) and the role of the social-cultural environments involved.

11. CULTURAL ROOTS OF LANGUAGE

Chimpanzees

Like many other mammals, chimpanzees have a number of more or less involuntary displays that express their mood, for example, piloerection indicating an aggressive mood, penile erection indicating a sexually receptive mood, and "play-face" indicating a playful mood. These evolved displays, and the skills of group members to read them and make correct judgements about their implications for behavior, are an important part of the social regulation of the group. In addition, however, chimpanzees also use a number of gestures intentionally. What marks these gestures as different from involuntary displays is that (a) they are clearly learned, because not all individuals use them (Tomase­ello, 1990); (b) they are used flexibly, both in the sense that a single gesture may be used in different contexts, and also in the sense that different gestures may be used in the same context, often in rapid succession when an initial gesture does not lead to the desired response (Plooij, 1978); (c) the initiator often waits expectantly for a response from the recipient after the gesture has been produced (Tomase­ello, George, Kruger, Farrar, & E. Evans, 1985); and (d) particular gestures are chosen or adjusted based on the state of the recipient; for example, gestures requiring visual access are only used when the recipient is looking (Tomase­ello, Call, Nagell, R. Olguin, & Carpenter, 1994). Even though they are used flexibly in the pursuit of interactive goals, however, these intentional gestures do not seem to possess the two key social-cultural characteristics of linguistic symbols.

With regard to reference and predication, we must note first that chimpanzees in their natural habitats employ basically two types of intentional gesture. First, attractors are imperative gestures aimed at activating the actions of others toward the self. For example, a well-known behavior from the wild is the leaf-clipping of adult males, which serves to make a noise that attracts the interest of females to their sexual arousal (Nishida, 1980). Similarly, when captive youngsters want to initiate play, they attract a partner by slapping the ground, poking at a desired partner, or throwing things at a potential playmate (Tomase­ello, Gust, & Frost, 1989). Chimpanzee attractors are thus invariably used dyadically, to attract others to the self; not triadically, to direct the attention of others to some entity in the environment. Because of this somewhat limited function, attractors most often attain their specific communicative value from their combination with involuntary displays.
OF 280

For example, the desire to play is most often communicated by means of an attractor that serves to gain attention to a "play-face." The second type of intentional gesture is metonymic gestures. These are mostly "intention movements" that serve to initiate interactive sequences. For example, play hitting is an important part of the rough-and-tumble play of chimpanzees, and thus many (but not all) individuals come to use a stylized "arm-raise" to indicate that they are about to hit the other and thus initiate play. Like attractors, most if not all metonymic gestures are imperative in function and used dyadically; that is, they are used to request a behavior of others toward the self, not to draw the attention of others to something in the external environment.¹

Overall, for neither type of chimpanzee gesture are there any convincing observations of individuals attempting to direct the attention of others to outside objects simply for the goal of sharing attention: so-called "declarative" gestures (Gomez, Sarria, & Tamarit, 1993). Moreover, there are certainly no observations of chimpanzees attempting to direct the attention of others to outside entities in the typically human ways of pointing to them or holding them up to show others. Chimpanzees thus do not seem to be using their gestures in acts of external reference as this concept is traditionally defined.²

With regard to the other key characteristic of symbols, their reciprocal nature, the most direct evidence comes from the processes by which young chimpanzees learn their gestures. There are two main possibilities. On the one hand, young chimpanzees might learn their intentional gestures by imitating those of their groupmates. Imitative learning of this type, in which the learner does not just mimic body movements but also learns the functional significance of a gesture, would seem to require that the learner understand the communicative

¹The major exception is food-begging, which is still an imperative gesture used by an individual in order to induce the foodbearer to produce action toward that individual herself.

²The well-known observations of Cheney and Seyfarth (1990) on the "referential" quality of vervet monkey alarm call vocalizations do not invalidate these observations. The vervet calls may not be intentionally produced or controlled, and indeed recent observations show that the alarm calls of chickens indicate specific predators in precisely the way that led researchers to call vervet alarm calls referential in the first place (Evans, Evans, & Marler, 1993). Also, Plooij (1978) presented several observations of wild chimpanzees that he believed to be acts of external reference, but each of these may be interpreted in a number of different ways.

intention of the gesturer so that he or she can reproduce the gesture when she has a similar communicative intention. The other possibility is a process called ritualization (or conventionalization). For example, two chimpanzee youngsters might begin to play through play hitting. On some occasion, one individual raises its arm in getting ready to play hit the other individual, who, because of past experience, is able to anticipate the impending hit and so begin the rough-and-tumble play at that early point in the sequence. The initiator then takes note of the recipient's anticipation and connects the raising of his own arm with the beginning of play, thus coming to use his arm-raise in a stylized manner, with no attempt to actually hit, and waiting for a response from the recipient. The arm-raise is now used intentionally, in order to elicit the desired response.

Recent research clearly indicates that chimpanzee gestures of both types are learned through a process of ritualization, not imitation. Most important, in an 8-year longitudinal study involving 2 generations of chimpanzees within the same captive group, Tomasello, Savage-Rumbaugh, and Kruger (1993) found that (a) some users raised gestures that no other group member used, (b) some youngsters used gestures that had not been directed to them and that they had had little opportunity to observe, (c) chimpanzees raised only with peers developed many of the same gestures as those raised with adults, (d) individual variability in types of gestures used was equivalent both within and across generations of the group, and (e) there was much individual variability in the way specific gestures were executed. Also, in an ongoing investigation, my collaborators and I have introduced a novel gesture to an individual in this same study (through shaping, outside the observation of other group members), and we have yet to observe the imitative learning of this gesture by others, even though they have had both the opportunity and the incentive to do so.

What is important for current purposes is that ritualization is basically a kind of social shaping process in which each participant learns the effects of its behavior on the other's behavior, sometimes in a complex sequential pattern. If chimpanzees acquire their gestures solely by means of ritualization—as seems empirically to be the case—this might
suggest that they understand their gestures from the perspective of one role only; that is, they understand the effects of the gesture from either the initiator's or the recipient's perspective, depending on which role they played in the conventionalizing interactions. In this scenario, they would not understand their own gestures as reciprocal communicative symbols, but rather as acts that produce certain predictable reactions in others, and they would understand the gestures of others as simple indications of their impending behavior. My only evidence for this admittedly speculative account is that chimpanzees do not imitatively learn their gestures from one another, implying that they do not understand that something they comprehend when another produces it is also something they may produce themselves when they have a similar communicative intention. And the reverse may also be true: They may produce a gesture but not comprehend it. My prediction is that if a chimpanzee had directed to a gesture that she herself had previously ritualized in production, but not in comprehension, she would not comprehend that gesture.

The reason that chimpanzees' intentional gestures are not symbolic in either of these two ways is that chimpanzees do not understand the behavior of others intentionally. First, chimpanzees do not use their gestures to refer to outside entities, much less to predicate things about them, because they do not understand that other chimpanzees may have intentional relations to outside entities. They attract others through various acts, and they even know that others need to be looking at their gestures to produce the desired effects, but they do not know that their conspecifics may attend to—or want or believe or in any way intentionally relate to—something outside their face-to-face relation. Such knowledge is essential (as will become apparent in the analysis of human children's early symbols) for the use of symbols in truly referential and predicative acts. Similarly, chimpanzees do not imitatively learn their gestures, and therefore do not understand them reciprocally, for precisely the same reason. To imitatively learn a new gesture requires that the learner comprehend another's gesture, including the goal toward which it is used, and then reproduce that gesture when she or he has a similar communicative goal. A gesture acquired in this way would presumably be understood reciprocally from the outset (from the perspective of the sender and receiver simultaneously) because understanding of the other's perspective was required for the imitative learning in the first place. If chimpanzees were able to understand directly the communicative intentions of others, they would learn to produce some gestures imitatively—perhaps even on first observation as children seemingly do in the acquisition of many linguistic symbols.

Communicative behaviors that are not used to direct the attention of others to outside entities and are not understood reciprocally are best called something such as "signals," not "symbols" because individuals may produce them intentionally, but not understand them intentionally when others produce them. Thus, signals are understood as spurs to action, and not as an intentional device that could be adopted by the self if needed. It should be noted that this interpretation of chimpanzee gestural communication is fully consistent with recent experimental results on chimpanzee acquisition of tool use, showing that whereas chimpanzees do learn some things by observing others, they do not engage in imitative learning of a type that requires an understanding of the intentions of others (Nagell, R. Olguin, & Tomasello, 1993; Tomasello, Davis-Dasilva, Camak, & Bard 1987). This hypothesis concerning the role of understanding others intentionally in the genesis and use of symbols will be explored and elaborated further as we deal with children and enculturated chimpanzees.

Human Children

Human infants have a number of intentional gestures that are most likely ritualized and understood in the same way that chimpanzees ritualize and understand their gestures, including such things as the "hands-up" gesture as a request to be picked up. Although there is certainly room for disagreement about how children understand these gestures, many of their earliest gestures, before their first birthdays, are not accompanied by the kind of spontaneous gaze alternation and checking with the adult that is typically used as evidence that the child understands the gestures reciprocally and predicatively (Butterworth, 1991). Moreover, in many cases the same may be said of children's earliest attempts at language. Many of children's early "words" are very likely learned through ritualization as well, that is, ritualization in which the child mimics an adult sound (word), with the result that the adult responds in some predictable and interesting way. These so-called presymbolic forms are often characterized as being simply a part of an activity, not a
symbol standing for anything else in the activity (E. A. Bates, 1979). We may thus call them vocal signals.

Symbol Use. Children's early words differ from chimpanzee gestures and from their own prelinguistic gestures and vocal signals with respect to both of the key social-cultural characteristics of symbols. First, children's earliest linguistic symbols may be used referentially to designate an outside situation. The predicative use of symbols allows for even more specific forms of reference. For example, upon observing a girl swimming in a pond, a nascent language learner might say to an adult something like "swimming," or "girl," or "water," or "wet," or "raft," or any number of other things. What is crucial in such cases is, first, that the child knows already that she and the adult are jointly attending to the event, and, second, that the child has a choice of which aspect of the event to talk about or comment upon: object, activity, property, location, or whatever. What the child is doing in such cases is choosing a particular psychological orientation or attitude from among other possible orientations and attitudes that he or she wishes the adult to take toward some entity for which they already share attention. The child is predicating something about a shared experience.

Although predication is most often thought of as a linguistically expressed comment on a linguistically expressed topic, a one-word predication is also possible if the child has first secured the adult's attention to a topic nonlinguistically (Tomasello, 1988). When a child holds up a shirt and says, "wet," that is an act of predication if indeed the child has other symbols to use in designating other aspects of the situation if he or she so chooses. Use of a symbol in this way implies that the child understands that the symbol is one among a number of options for directing the attention of the other person to one aspect, as opposed to other possible aspects, of the shared situation. Young children engage in one-word predication defined in this way from before the middle of the second year of life.

Symbol Learning. Excluding cases in which vocal signals are transformed into symbols, children learn their early words by some form of imitative learning from adults. It is thus presumably the case that they understand them reciprocally. This means that when children predicate something of a situation to an adult, they, in effect, understand or anticipate the effect of that predication on the adult's construal of the situation. It also means that they would understand another individual's use of that same symbol. This reciprocity is a natural outcome of the imitative learning process because imitative learning requires that the child understand, not only the adult's communicative intention in using a novel piece of language, but also the fact that he or she may use the same piece of language when she has the same communicative intention.

To understand more precisely how imitative learning leads to the reciprocal understanding of symbols and their use as predicative devices, several recent studies concerning children's comprehension of adults' referential intentions are instructive. First, Baldwin (1991, 1993) had an adult look at one object, while a young child was looking at another object, and then say "It's a toma." In this situation, children did not suppose that the adult was naming the object they were seeing, but rather the object at which the adult was looking. Children were not passively associating a sound and a perceived object. They were actively inferring adult referential intentions. Along similar lines, Tomasello and M. Barton (1995) had an adult announce to 24-month-old children that he or she was going to go find a "toma" in a row of buckets. The adult then approached the buckets, took out two objects and frowned at them, then picked up a third object with obvious glee. Almost all of the children understood immediately that the adult intended for the word "toma" to indicate the third, not the first or second, object chosen. Again, children were not passive participants, but were actively monitoring the adult's intentions (via facial expression and other behavioral cues) to determine the intended referent. Finally, Tomasello and M. Barton (1995) also had an adult announce his or her intention to "plunk" an object, and then proceed to perform one action accidentally followed by another purposefully. Young children understood that the new symbol was intended to depict the action the adult performed intentionally, even though the accidental action was the one performed immediately after the new word was said.

The point is this: To understand novel words in cases such as these, young children use a variety of social-pragmatic cues to determine precisely what the adult is doing and why—that intention was behind the action. Children must understand that the adult using the new word wants them to focus their attention on one specific aspect of the upcoming event out of many aspects that might be possible referents. When the child then uses the word, with the same communicative
intention as the adult had in the original learning situation, the reciprocity is apparent. Thus, understanding the exhortations of others that one should attend to a situation in a certain way, and adopting them for oneself in exhorting others to take similar orientations in similar contexts, is called by Tomasello, Kruger, and H. H. Ratner (1993) "imitative learning." It is considered one form—the ontogenetically first form—of a more general process called "cultural learning." The defining feature of cultural learning is its direct dependence on children's ability to understand the intentions of another person who attempts both to direct the children's attention to outside entities and to comprehend their efforts to do the same thing in return. What is unique about language acquisition is that the preexisting set of linguistic symbols provided by the culture offers to adult and child a common medium by means of which they may engage in this form of interaction reciprocally, using a common set of devices to encourage one another to take particular psychological orientations in shared situations. Thus, the fact that young children can imitatively learn novel words in situations in which there are many possible targets for adults' referential intentions indicates that from a very early age their understanding and use of symbols are both reciprocal and preceptive.

My explanation for the unique qualities of human children's early linguistic symbols is that children, unlike chimpanzees in their natural habitats, perceive and understand the behavior of others intentionally. Because they understand the behavior of others intentionally, they can also comprehend their communicative intentions in using particular symbols and thus imitatively learn that use for themselves when they have similar communicative intentions. This makes the symbols reciprocal. Similarly, understanding that other persons may intentionally direct attention to various aspects of a shared situation allows children to comprehend and produce linguistic symbols with those persons in a preceptive manner. In all, understanding other individuals as intentional agents is the key social–cognitive skill that accounts for the difference between signals and symbols.

It should be mentioned, at least in passing, that further evidence for the view that young children's understanding of intentions is crucial to their acquisition of symbols comes from the emergence of a number of new behaviors at around the same age when young children are learning their first words. Thus, at around their first birthdays human infants for the first time systematically look where others are looking, adopt their emotional attitude toward situations (social referencing), and imitatively learn their instrumental behaviors toward objects. These all require that the young child understand the adult as an intentional agent whose attention or behavior to outside entities may be followed and adopted. Thus the stage is set for language acquisition (Tomasello, 1995). Perhaps even most interesting in the current context, as children begin to learn their first symbols, they also begin to use their gestures in a more interpersonally sophisticated way, often checking with adults to see if they have followed a pointing gesture, for example. Some children even learn a new type of gesture at this age—what has been called the "symbolic gesture" (Acredolo & Goodwin, 1993)—involving such things as flapping the arms to a bird, or blowing toward liquids that are hot. These are presumably learned imitatively and understood reciprocally. It is also important that Petitot (1992) found a distinct difference in the way young deaf children point prelinguistically and the way they point when they begin to do so symbolically in American Sign Language—again at around this same age and also relying, presumably, on imitative learning and reciprocal understanding. The important point is that all of these behaviors evidence the 1-year-old child's emerging ability to understand others as agents with intentional relations to various aspects of the world. The behaviors thus provide further support for the hypothesis that the acquisition of linguistic symbols as reciprocal communicative devices that can be used in acts of predication depends crucially on the understanding of others as intentional agents.

The Cultural Context. Obviously, to learn the use of a new linguistic symbol in its conventional communicative context children must be exposed to others using that symbol. The cultural environment is thus in some sense presupposed by the notion that children acquire their words by means of cultural (imitative) learning. In addition, however, it may be that some more specific aspects of the cultural environment are required. Thus, recent research has demonstrated that children learn almost all of their earliest language in cultural routines or in joint attentional interactions of one sort or another. Joint attentional routines such as feeding, diaper changing, bathing, interactive games, book reading, car trips, and a whole host of other activities constitute the formats within which children acquire their earliest linguistic symbols (see Bruner, 1983, and Nelson, 1985, for summaries). These routines scaffold the initial language acquisition of children in the sense that they
create, with no need of a conventional language whatsoever, a shared referential context within which the language of the adult makes sense to prelinguistic children—assuming their ability to understand the behavior of others intentionally.

Even after they have learned their first words, children continue to acquire the vast majority of their language inside nonlinguistically understood joint attentional episodes, and individual differences in the rate at which children acquire their early words are predicted by the extent to which they participate with adults in joint attentional episodes (see Tomasello, 1988, for a review). Again, such episodes, whether or not they are routine, require both a certain type of social understanding on the child's part and certain kinds of participation in the interaction on the adult's part. The nature of routines and joint attentional episodes may vary across cultures to some extent, but there are also certain commonalities in the structuring of routine activities of very different content across different cultures, and it is likely that these commonalities play some role in the early language acquisition of all children (Peters & Boggs, 1986).

Thus, the overall point is that human beings come into the world adapted for intentional communicative interactions with others and that, not surprisingly, cultures are structured in ways that facilitate this process. Although research shows that children learn much of their early language inside routinized cultural interactions, it is still possible to argue from this research that such interactions are not entirely necessary, but only useful to speed up the process somewhat. I would argue for their necessity, however, by insisting not only that routine interactions are necessary for children's early language acquisition—because only in such contexts can their still fragile social–cognitive skills to understand the intentions of others operate—but that children need many months of social interaction with others before language acquisition begins if they are to come to understand others at all as intentional agents (and thus to actively participate at all in cultural routines). Following the arguments of Kaye (1982), I would argue that social interactions of a certain type (i.e., interactions in which infants are themselves treated intentionally) are necessary in children coming to understand other persons as intentional agents. There is, of course, very little direct evidence for this position in research with human children (inferences from wild children and neglected children are notoriously problematic), but research with chimpanzees exposed to human-like forms of symbolic communication provides some highly suggestive findings.

Enculturated Chimpanzees

The first attempts to teach chimpanzees skills of symbolic communication relied on various behavioristic training techniques. Experimenter would do such things as hold up objects, shape chimpanzee hands into signs, and reinforce the connection (production training); or they would themselves give a sign and reinforce chimpanzees for selecting the corresponding object (comprehension training). Subjects trained in such ways did learn to do some very interesting things communicatively, but what they learned may not have been true symbols according to the current definition (Savage-Rumbaugh, 1986). Thus, it turns out that when these pioneer chimpanzees were taught symbols in production, they could not comprehend them, and when they were taught to comprehend them, they could not produce them appropriately (Savage-Rumbaugh, 1990). Their symbols were either production based or comprehension based, depending on how they had been trained. They were not understood reciprocally in the sense that they could mediate communicative interactions in both directions: from chimpanzee to human or from human to chimpanzee.

In contrast, the pygmy chimpanzee (bonobo) Kanzi (and a common chimpanzee since) was raised more naturally in a human-like cultural environment with linguistic symbols appearing as a natural component of his social interactions with humans; he was not conditioned or trained in the use of human symbols in any way. Kanzi was regularly invited into highly structured cultural routines such as changing diapers, preparing food, going outdoors, taking a bath, blowing bubbles, riding in the car, looking at a book, and playing various games. Thus, language (both spoken English and a lexicogram board displaying abstract symbols) was used by his human caretakers in these situations in any way that seemed natural to them, presumably in ways comparable to those by which human children are exposed to language. As a consequence of this more culturally saturated and less behavioristic upbringing, Kanzi learned to comprehend and use symbols reciprocally in very human-like ways. Because his productive behaviors were not shaped through reinforcement, Kanzi learned first to comprehend linguistic
symbols—as human children do—and only then, with no specific training, did he begin to produce them spontaneously (Savage-Rumbaugh et al., 1993). This initial learning through observation and imitation thus ensured, as it does for human children, that the symbols produced were understood reciprocally.

Savage-Rumbaugh's work would seem to be a very convincing demonstration that routine cultural activities and events help to structure the language learner's experience in a way that is essential to the acquisition of communicative symbols. However, the effect is actually more general than this. Two recent studies compared other social-cognitive and cultural learning skills that chimpanzees develop in cultural environments as contrasted with those that they develop in more typical captive environments. For both imitative learning (Tomasello, Savage-Rumbaugh, & Kruger, 1993) and joint attention (Carpenter, Tomasello, & Savage-Rumbaugh, 1995) enculturated chimpanzees were found to be more similar to 2-year-old human children than they were to their conspecifics who were not raised in cultural environments. Documentation of humanlike behaviors in all three of these domains—language, imitative learning, and joint attention—none of which were the result of explicit training, argues very strongly that through interacting with other persons inside natural cultural routines chimpanzees are not just learning some isolated skill, but rather are coming to learn more generally about the behavior of human beings. They are coming to view humans as intentional agents, and this allows them to comprehend and then to produce appropriately the symbols humans use in their interactions with them, as well as to engage in a number of other human-like cultural skills.

None of this is meant to imply, of course, that cultural environments create social-cognitive skills de novo. Many animals are raised in humanlike ways and do not develop such skills. The limitations of a cultural environment are even more dramatically demonstrated by the case of autistic children who have been raised in all kinds of cultural environments, with and without specific intervention programs, and who still do not develop any language or imitative learning skills. (This describes about half of all autistic individuals.) It is also important in this context to note that recent research has established that the extent to which autistic children can engage in nonlinguistic joint attentional interactions with adults—implying some understanding of adults as intentional beings—is strongly correlated with extent of their symbolic and linguistic skills (Landry & Loveland, 1986; Mundy, Sigman, & Kasari, 1990).

Lexical Symbols: Summary

Most recent accounts of early linguistic symbols have focused on processes of information representation and decontextualization as their defining features (E. A. Bates, 1979; Huttenlocher & Higgins, 1978). Although not disputing that these are an important part of the process of learning to understand and use symbols, I simply do not think they are the whole story, or even the most important part of the story. In my view the kinds of social understandings discussed here—the reciprocal understanding of symbols along with the understanding that linguistic symbols allow choices in the particular orientation or attitude that one might predicate of a situation—are crucial in distinguishing symbols from other kinds of semiotic devices. In the current hypothesis, the understanding of symbols as both reciprocal and as predicated is made possible by the child's developing understanding of other persons as intentional agents. Such understanding allows children to comprehend that when the adult uses a symbol to refer to some aspect of the current situation, the adult intending that the child attend to that aspect. Linguistic symbols also require certain forms of social-cultural interaction with others. Individual organisms do not just come to invent things such as symbols on their own. They must interact with others inside contexts that make apparent the various intentions involved, it is essential that they also have a period of time in which other persons treat them intentionally before they can understand other persons as intentional agents at all.

Grammatical Symbols

As in the case of symbols, there is a long history of controversy over the proper characterization of grammar. Generative grammarians, for example, have defined grammar and syntax in terms of formal mathematical rules in which semantic and communicative functions play absolutely no role (Chomsky, 1986). The cultural roots of such rules are presumably nil. But the generative view of grammar is not the only
possible view. The view of grammar emerging in cognitive linguistics relies not on formal rules but explicitly and exclusively on symbols and categories of symbols (Langacker, 1987). Competence with a language is thus nothing more or less than competence in using symbols, or, more precisely, in using a "structured inventory of symbolic devices," of which there are many types serving many different functions, including grammatical functions (see Wittgenstein’s analogy of a tool box, 1953). To construct communicative messages, human beings take symbols and categories of symbols from this inventory and integrate them into larger symbolic wholes that relate in comprehensible ways to the ongoing communicative context.

For current purposes, the cognitive linguistics view of grammar allows us to take our previous analysis of symbols and use it as a foundation for our characterization of grammar. The basic idea is that human beings, including young children, live in a world of events and actions (Nelson, 1985). As they begin acquiring their first symbols, they learn to understand and to predicate symbolic references to different aspects of the same event: the entities, relations, and processes that constitute it. The preexisting language that children encounter in their interactions with adults provides a number of different means for these predictions and for coordinating these predications into one coherent description in which the roles of different participants in an event are symbolically marked with grammatical symbols (i.e., into "sentences"). Thus, what we might call lexical symbols designate the elements in a complex cognitive scene, whereas grammatical symbols operate on these elements to designate the structure of the scene—the relation of the elements to one another in the event (Talmy, 1988). The nature of these grammatical symbols is different in varied languages. For example, in some languages speakers mark the agent and recipient of an action by placing them in particular orders. In other languages the order of the elements is not crucial, but there are special case endings. In still other languages there are special symbols (such as some English prepositions) or special intonation patterns (as in tone languages) that serve these same functions (E. A. Bates & MacWhinney, 1989). The point is that grammatical symbols are simply another kind of symbol in the communicative inventory, created and used for purposes of communicating complex messages concerning recurrent and predictable relationships among entities in the shared world of their users.

11. CULTURAL ROOTS OF LANGUAGE

It is clear from this characterization that the most important cultural roots of grammar are precisely the same as those of individual symbols because grammatical symbols are just another form of linguistic symbols. In addition, however, the combining of individual symbols and the marking of their respective roles with grammatical symbols entails some further cognitive and social–cognitive skills that are cultural in ways that go beyond individual symbols. Again, the best way to explore this issue is by comparing the grammatical skills of chimpanzees, children, and enculturated chimpanzees.

The ability to take an event, parse it into elements, relations, and processes of various types, and then to indicate these elements, relations, and processes with individual lexical and grammatical symbols in one coherent utterance, is an ability that has evolved only in the human species. Chimpanzees in their natural habitats sometimes string multiple vocalizations or gestures together, but they do not use these combinations to create new meanings of any type or to parse (predicate) events into their constituent elements. Most chimpanzee gesture combinations are repetitions of the same gesture or several gestures, all with the same communicative intention (e.g., to request play); they do not contain different gestures playing different roles within a single communicative intention (Tomasel10, Savage-Rumbaugh, & Kruger, 1993). Chimpanzee gestural communication thus cannot display a grammar because their gestures are not referential and predicative in the first place.

Human Children

Early Grammatical Skills. Children’s grammatical skills begin with their skills of predication. In their very first multilword utterances in their second year of life, human children predicate different aspects of the same event, saying for example, "More juice," and on another occasion of the same event, "Mommy juice." These combinations usually follow closely on the heels of their first one-word predications. In these early word combinations children are not employing grammatical symbols to designate the relation between the lexical symbols, however. They utter the words in a certain order, but they are not actively using that order to designate the roles that individual words are playing: "More juice" does not mean something different from "Juice
more," and the children are not explicitly marking the intended relation in any other way. In these early utterances it is up to the listener to infer what grammatical relation among lexical symbols the child is intending. In many cases children follow the adult word order and therefore seem to be using adult-like grammatical marking, but I am aware of no evidence showing that children's early word combinations employ word order or any other grammatical symbols as productive communicative devices.

When children do begin to actively mark lexical symbols for their grammatical functions, they do so in event-specific ways, in which event most often means intentional human action. Thus, children learn first grammatical symbols inside specific cultural activities and events, most typically activities and events about which they have previously predicated different things (Ninio, in press). For example, English-speaking children learn the grammatical symbols for "hitting" in situations such as sibling disputes while eating at the dinner table, swimming at the pool, and deciding who will play with certain toys. What they learn is tied to the specific event and the way adults talk about the intentional actions within that event, for example, that "the hitter" goes before "hit" and "the thing hit" goes after, whereas the "thing hit with" is designated by the word "with." Evidence for this view comes from the diary study of Tomasello (1992a) who found that his subject did not generalize the use of grammatical symbols across different event structures: She had to learn how to parse and grammatically mark the components of each of the different events individually. R. Olguin and Tomasello (1993) provided experimental confirmation for this explanation: the so-called Verb Island Hypothesis. These results mean that the syntagmatic categories of children's early language are event-specific categories tied to the specific participant roles involved, such as "hitter," "thing hit," "kisser," and "thing kissed," not event-general participant roles such as "agent," "patient," and "instrument." It is also interesting and important for current purposes to observe that almost all of these early events serving to organize children's early grammars are the intentional actions of persons, including the child in almost all of

---

1. It should be noted that children of this age can substitute different objects in a single participant role without having heard that object's name used in that role (e.g., "More juice," "More dax"), providing for some degree of productivity and indicating some early notion of something like a grammatical category of noun (Tomasello & Olguin, 1993).

2. The major exceptions to this generalization were rational words or verbs used to indicate the locations or simple spatial transformations of objects.
least some subset of events, such as transitive actions). Formation of
this paradigmatic category makes possible more general such syntag-
matic categories as "agent" and "patient" as generalized participant
roles across events. With the category of verb in hand, children may
now understand that the statement, "John is glorping Mary," is intended
to indicate that John is doing something to Mary and not the reverse
(L. R. Gleitman, 1990). We may now speak of the beginnings of a lan-
guage system, based on fundamental processes of linguistic categoriza-
tion, by means of which the child may create and comprehend an in-
finite variety of linguistic expressions.

The process of linguistic category formation is fundamentally a
cognitive process. However, as my central concern in this chapter is
with the cultural roots of language acquisition, I would like to point out
two ways in which it is also social-cultural. First, in all of the world's
languages that have been studied there are certain commonalities in the
complex events and participants that young children are first motivated
to verbalize. Many of these commonalities are captured in what Slobin
(1985) called the "Manipulative Activity Scene": situations in which
people do things to objects, often in particular settings and sometimes
with particular instruments—such things as giving, bringing, opening,
throwing, and a variety of other concrete activities that humans regu-
larly perform on objects and with other people. These actions typically
have animate actors (sometimes with instruments) and inanimate re-
cipients of the action (sometimes in a particular setting or with another
person). As Bruner (1990) pointed out, this basic schema is very simi-
lar to the basic structure of all kinds of cultural narratives in which an
actor acts to effect an outcome in some setting with some instrument.
What this means is that the most basic participant roles underlying chil-
dren's earliest grammatical creations—such things as agent, action, pa-
tient, instrument, and location—are in reality the very same categories
in terms of which they understand the most basic cultural activities. In
both language and social-cultural activities in general, children, like
their cultures, are concerned with the same basic types of participant
roles.

The second way that early grammatical categories are social is
somewhat indirect. An important part of early grammar structure is the
formation of paradigmatic word classes such as nouns and verbs that li-
cense much productivity; a new word identified as a noun (e.g., by the
presence of an article) may be used in many appropriate ways in which

11. CULTURAL ROOTS OF LANGUAGE

it has never been heard. In the analysis of Tomasello (1992a) the con-
struction of these categories is based on the subject's reflection on his or
her own linguistic productions. This is the same process as that pro-
posed by Karmiloff-Smith (1992), who found that in many cognitive
domains children first learn individual procedures (such as lexically
specific structures), and then, through reflecting on their activities,
construct general principles (redescriptions) on which to base their fu-
ture behavior. Paradigmatic categories can thus be formed only when
the child has reflected on the way he or she has used various concrete
linguistic items, just as principles of lexical acquisition can only be
formed after the child has learned to use some individual lexical items.

But self-reflection, in the hypothesis of Tomasello, Kruger, and H. H.
Ratner (1993), is nothing other than cultural learning turned on the self
and its behavioral-cognitive products. That is to say, the ability to self-
reflect is simply the ability to take the point of view of another individu-
al, real or imagined, as they are observing me. Self-reflection is my
acting as if I were another person, looking at my behavior from the
outside, using all of my basic powers of human perception and catego-
rization in the process; I am culturally learning from, and thus forming
categories of, my own unreflective productions. Grammatical catego-
ries are thus formed through a variation on the basic processes that un-
derlie the acquisition of lexical and grammatical symbols.

It should also be noted that later in the language development
children begin to use a variety of specialized discourse structures that
differ in various ways from the prototypical events of interest to 2-year-
olds. Children learn to produce the conventional form of such things as
questions, passive sentences, and sentences with embedded clauses.
However, it turns out in these cases as well that children learn these
structures on an individual basis, relying, for example, on their knowl-
edge that certain kinds of events, by their very nature, structure other
events (e.g., think and believe require other events to complete them;
see L. Bloom 1991, on the use of specific "wh" words), or on the gram-
matical marking required when one wants to highlight the patient of an
action. The role of the social-cultural environment in children's acquisi-
tion of these more complex forms of grammatical competence is an
issue that has not received much attention. It seems clear, to me at
least, that children must be exposed to these structures in situations in
which they have some way of understanding the communicative inten-
tions of the speaker and connecting them with the new form they are
hearing (Akhtar, 1995). Evidence that this is indeed the case comes from studies such as those of Nelson (1977) and Farrar (1990), who both found that adults' replying to nonadult-like child utterances by recasting them into adult form facilitates children's acquisition of some types of grammatical structures, presumably because these recastings represent the child's own semantic intentions. Also, E. A. Bates and MacWhinney (1989) have shown how efforts to take the perspective of the listener in discourse (as manifest in topic-comment structure) underlie the basic grammatical category of sentence subject. Also Rispoli (1995) has shown in an even more detailed way how various grammatical structures derive from basic patterns of discourse-communicative interactions. Learning grammatical symbols and structures is essentially the same process as learning individual symbols.

Enculturated Chimpanzees

The story of chimpanzees enculturated in the use of grammar is very similar to the story of their enculturation in the use of individual symbols, once again highlighting the role of the cultural environment. Some early attempts at training chimpanzees in grammar did not work very well. The trained chimpanzees produced "words" and strings of words as they were reinforced for doing so, but for a number of reasons these words did not seem to be grammatical in the sense that they did not parse events and designate symbolically the different participants in the events. Thus, Terrace, Petitto, Sanders, and Bever (1979) showed that many of their subject's word strings were repetitive, almost random, and did not add significantly to the description of the event as more words were added. Even in recognizing such utterances as Washoe's famous "water bird" as creative symbol combinations, it is still very likely that Washoe was not employing word order in this utterance as a productive grammatical symbol; it is most likely analogous to children's first word combinations, which leave it up to the listener to infer the intended relation. Of course, the failure of trained chimpanzees in the domain of grammar is not surprising if indeed their "words" were not symbolic, if they were not reciprocal symbols used predicatively in the first place.

The most convincing evidence for an understanding of grammatical symbols by an ape comes once again from Kanzi. Kanzi clearly under-
Grammatical Symbols: Summary

If we regard grammar as just one of the devices present in human languages for conceptualizing and communicating shared cultural experiences, then we may begin to study the psychological and cultural processes involved in grammatical development. I argue that the development of grammatical competence in both human children and enculturated chimpanzees is cultural in at least three senses. First, the kinds of things they talk about are recurrent cultural activities, parsed into the actors, actions, affected object, instrument, and setting information characteristic of cultural narratives of all types. Moreover, at an early stage for children, and perhaps at all ages for enculturated chimpanzees, these activities serve as the major organizing structures in early grammars. Second, the parsing of events relies on an understanding of lexical symbols as predicative, which, in turn, relies on an understanding of other persons as intentional. Thus, the social–cognitive skills that children use to acquire grammatical symbols are precisely those that have evolved for the acquisition of cultural skills in general: abilities to enter into a joint attentional interaction with intentional agents and to take their perspective on the situation. The ability to reflect on the structures constructed in this way may be essential to grammatical category formation and may derive from cultural learning abilities in general. Finally, the kinds of discourse-communicative functions that arise in linguistic interactions with others are an integral part of the acquisition of the more complex sorts of grammatical competence: question asking, passive sentences, and the like (which may be the exclusive province of human beings). In all, looking at grammar as an extension of human symbolizing skills allows us to view lexical and grammatical symbols within the same cognitive and communicative contexts, and to compare children of different ages and even nonhuman species as they attempt to acquire human adult-like grammatical skills.

THE EVOLUTION OF LANGUAGE

There is a growing consensus among behavioral biologists that what is unique about primate intelligence is the nature of its adaptations to the social environment, and recent research has elucidated many of the ways that primates may understand, predict, and influence the behavior of conspecifics (Byrne & Whiten, 1988; Cheney & Seyfarth, 1990). In this context human beings have developed some of their own unique, species-typical skills of social cognition and social learning that make possible the process by which one generation of human beings assimilates the cultural knowledge and skills of the generation preceding them (Tomasello, Kruger, & H. H. Ratner, 1993). It is the central contention of this paper that human languages have their roots in general cultural skills of this type, and that they are transmitted across generations in the same way as other cultural skills. The kinds of abstract linguistic structures that theorists such as Pinker and P. Bloom (1991) claimed are not transmittable in this way are simply not a part of the structure of language as conceived by a growing group of linguists and psycholinguists, including myself, who reject the mathematical approach to language structure.

Because of their mathematical construal of linguistic structure, generative grammarians such as Chomsky and Pinker cannot really say anything useful about language evolution except that it burst onto the scene rather recently. I think it is important in this regard, though, to remember that individual lexical symbols are uniquely human inventions as well. The generative view of language must then posit two evolutionary salutations if it is to account for the totality of human grammatical competence: one for symbols and another for grammar. Moreover, given their view of grammar, it would seem that generative grammarians have to posit very different kinds of adaptations in the two cases: one psychological and the other formal. Much more plausible, I think, is the view that skills of human communication emerged gradually from skills of primate social cognition and communication, and that as those underlying skills changed, so did the nature of the human communicative system. Thus, intentional gestures of the type used by modern chimpanzees may have evolved at some point in human evolution into something like linguistic symbols as human beings came to understand the intentions and perspective of other human beings more clearly, which then led to the kinds of reciprocal understandings and predications characteristic of 2-year-old children's linguistic symbols. Grammatical symbols may have evolved from this base as humans came to understand that acts of predication made sequentially may be bundled into one coherent utterance in which different aspects of an event are symbolized by different lexical and grammatical symbols.
Obviously, I do not know the specific evolutionary steps and timetable for all of this. An interesting twist in the story, however, is provided by the chimpanzees raised in human-like cultural environments. Our experience with these apes must lead us to ask the question of whether there might be specific types of social interactions—types practiced by all human cultures of the world—that are necessary for the language acquisition of children to proceed normally. Although we do not know for sure, it seems evident by comparing the different procedures used with different degrees of success with human-raised apes that the key element may be the learner's participation in routine cultural activities in which an adult human treats the ape intentionally by directing its attention, encouraging its behavior (including imitation), and the like. Chimpanzees are not capable of creating among themselves these kinds of cultural activities, but they are able participate in and take advantage of such activities. This makes plausible the view that human skills of social cognition and cultural learning have co-evolved with the cultural environment and the way it is structured. It is thus possible that there are kinds of environments—with no people or in which people behave in unpredictable ways—in which human children would not acquire any linguistic skills at all.

To conclude, I will say only that the "scorched earth" policy used by generative grammarians to keep psychologically oriented linguists at bay is working no longer (E. A. Bates, 1984). Cognitive and functionally oriented linguists no longer believe that it is their duty to explain the kinds of mathematical structures that generative grammarians invent with numbing speed and regularity. Before we can discover the psychological bases of human linguistic communication skills, it is necessary first to describe the different structures of the world's various languages in ways that make their connection with human cognition and communication discoverable. Now that such descriptions are becoming available, it is possible for the first time for psychologists to make these connections in ways that address the kinds of concerns that motivate their discipline. I have proposed here, in a preliminary fashion, some ways that human language acquisition and use depend on, indeed are part and parcel of, the kinds of cultural interactions and understandings that are the reality of human children in their second year of life and beyond.

ACKNOWLEDGMENTS

I thank Nameera Akhtar, Josep Call, Malinda Carpenter, Kathy Nagell, Jennifer Ashley, Andrew Rosner, and Boris Velichkovsky for their helpful comments on an earlier version of the manuscript.

REFERENCES


11. CULTURAL ROOTS OF LANGUAGE


11. CULTURAL ROOTS OF LANGUAGE


