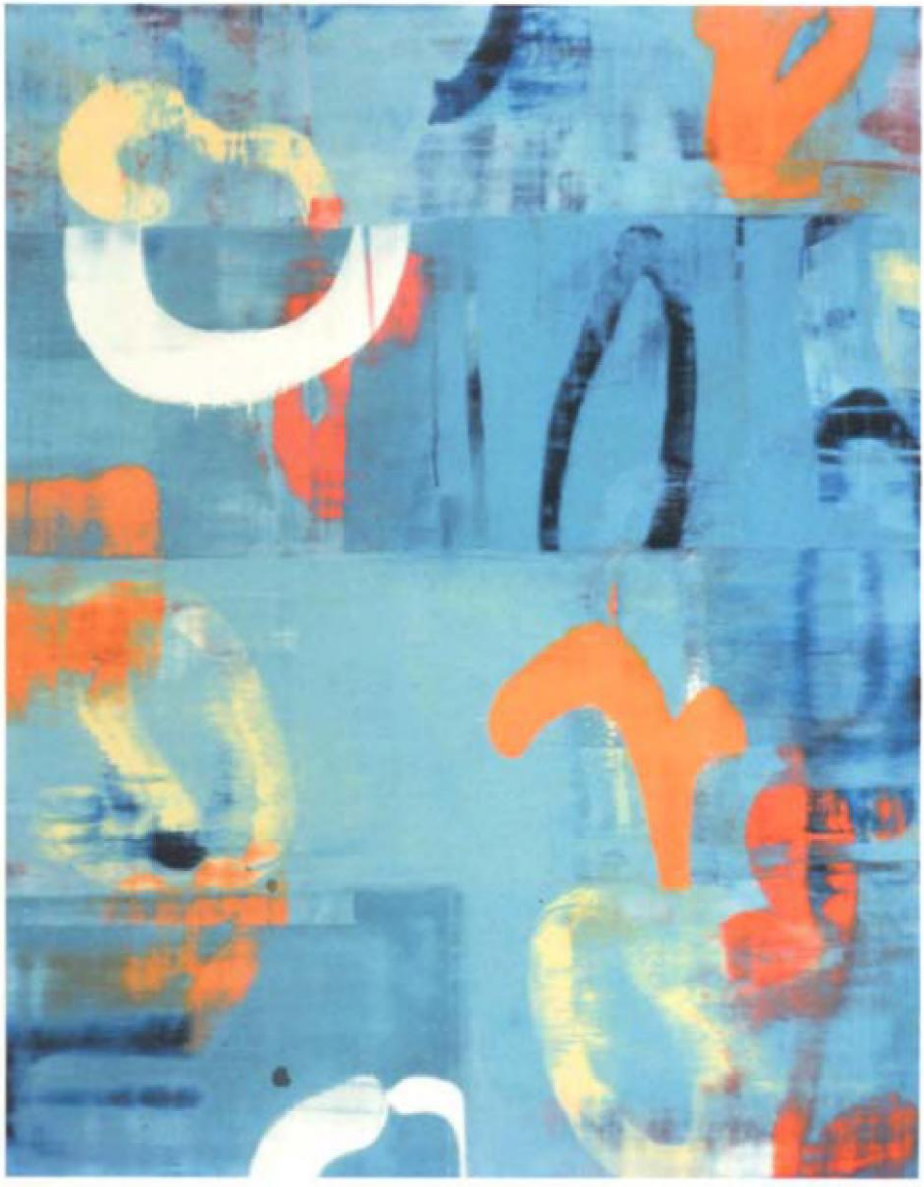
Michael Tomasello

CONSTRUCTING A LANGUAGE

A Usage-Based Theory of Language Acquisition



Printed in the United States of America
First Harvard University Press paperback edition, 2005

Library of Congress Cataloging-in-Publication Data

Tomasello, Michael.
Constructing a language : a usage-based theory of
language acquisition / Michael Tomasello.
p. cm.
Includes bibliographical references and index.
ISBN 0-674-01030-2 (cloth)
ISBN 0-674-01764-1 (pbk.)
1. Language acquisition. 2. Cognition in children I. Title.

P118 .T558 2003
4019.93—dc21 2002038840

Designed by Gwen Nefsky Frankfeldt

Contents

1 Usage-Based Linguistics 1

2 Origins of Language 8

2.1. Phylogenetic Origins *9*
2.2. Ontogenetic Origins *19*
2.3. Children's First Utterances *31*
2.4. Summary *40*

3 Words 43

3.1. Early Words and their Uses *44*
3.2. Processes of Word Learning *58*
3.3. Theories of Word Learning *81*
3.4. Summary *91*

4 Early Syntactic Constructions 94

4.1. The Nature of Constructions *98*
4.2. Early Constructional Islands *113*
4.3. Marking Syntactic Roles *126*
4.4. Summary *140*

5 Abstract Syntactic Constructions 144

5.1. Abstract Constructions *146*
5.2. Constructing Constructions *161*
5.3. Constraining Constructions *175*

5.4. Theories of Syntactic Development *181*
5.5. Summary *193*

6 Nominal and Clausal Constructions 196

6.1. Reference and Nominals *199*
6.2. Predication and Clauses *213*
6.3. Learning Morphology *232*
6.4. Summary *241*

7 Complex Constructions and Discourse 243

7.1. Complex Constructions *244*
7.2. Conversation and Narrative *266*
7.3. Summary *279*

8 Biological, Cultural, and Ontogenetic Processes 282

8.1. Dual Inheritance *283*
8.2. Psycholinguistic Processes of Acquisition *295*
8.3. Psycholinguistic Processes of Production *305*
8.4. The Development of Linguistic Representations *314*
8.5. Summary *320*

9 Toward a Psychology of Language Acquisition 323

References *331*
Acknowledgments *373*
Index *375*

Constructing a Language

ONE

Usage-Based Linguistics

The confusions which occupy us arise when language is like an engine idling, not when it is doing its work.

—LUDWIG WITTGENSTEIN

NOTHING could seem less remarkable than a one-year-old child requesting "More juice" or commenting "Doggie gone." But the remarkable fact is that even these baby utterances differ from the communicative activities of other animal species in a number of fundamental ways. For example, other animals do not refer one another's attention to outside entities such as juice, they do not make disinterested comments to one another about missing doggies or the like, and they do not combine communicatively significant elements to create new meanings. But from an ethological perspective, perhaps the most astounding fact is that something on the order of 80 percent of all *Homo sapiens* cannot understand these simple utterances at all. That is, whereas the individuals of all nonhuman species can communicate effectively with all of their conspecifics, human beings can communicate effectively only with other persons who have grown up in their same linguistic community—typically, in the same geographical region.

Whatever may be the evolutionary reasons for this unique, indeed bizarre, situation, one immediate outcome is that, unlike most other animal species, human beings cannot be born with any specific set of communicative behaviors. Young children must learn during their individual ontogenies the set of linguistic conventions used by those around them, which for any given language consists of tens of thousands, or perhaps even hundreds of thousands, of individual words, expressions, and constructions. The human species is biologically prepared for this prodigious task in ways that individuals of other species are not, of course, but this preparation cannot be too specific, as human children must be flexible

enough to learn not only all of the different words and conventional expressions of any language but also all of the different types of abstract constructional patterns that these languages have grammaticized historically. It thus takes many years of daily interaction with mature language users for children to attain adult-like skills, which is a longer period of learning with more things to be learned—by many orders of magnitude—than is required of any other species on the planet.

The first proposal in the modern context was that young children learn their "verbal behavior" using the same garden-variety learning mechanisms they use to learn other behaviors—which, by the way, are the same learning mechanisms used by rats and pigeons. Thus, Skinner (1957) proposed that young children learn pieces of language by means of instrumental conditioning (based on principles of association) and that they generalize to new instances by means of stimulus generalization (based on principles of induction). But in his withering review of Skinner's book, Chomsky (1959) argued that there are some principles of grammar that are so abstract and, in a sense, arbitrary that children could not possibly learn them by means of simple association and induction. Indeed, Chomsky (1968, 1980a, 1986) later argued that there are some abstract principles of grammar for which children have no reliable and unambiguous evidence at all—given that the language they experience consists of nothing more than a series of individual utterances. This is the so-called argument from the poverty of the stimulus, and Chomsky's well-known solution was to hypothesize that human beings are born with an innate universal grammar containing a number of abstract principles that guide the acquisition process.

This argument had a profound effect on researchers studying children's language in the 1960s and 1970s. The prevailing opinion at the time was that baby utterances such as "More juice" and "Doggie gone" were just that, baby utterances that rested on very concrete and seemingly non-adult-like linguistic representations such as *More X* and *X gone* (e.g., Braine, 1963, 1976). But people impressed with the argument from the poverty of the stimulus looked at these baby representations and at the formal descriptions of adult language being proposed by Chomsky and others and said, in effect: "You can't get there from here" (e.g., Gleitman and Wanner, 1982). The majority opinion in the field thus changed rather quickly to the view that children's early language was somehow undergirded by some kind of linguistic abstractions—perhaps even the same ones that underlie mature adult language. This is the so-called *continuity assumption:* that basic linguistic representations are the same throughout all stages of child language development—since they come ultimately from a single universal grammar (Pinker, 1984).

But much has happened in the last two decades in developmental psychology, linguistics, and cognitive science which suggests a re-evaluation of the situation, that is, which suggests that children *can* get from here to there, and that they can do it without the aid of any hypothesized universal grammar. There are two fundamental points: (1) children have at their disposal much more powerful learning mechanisms than simple association and blind induction; and (2) there exist plausible and rigorous theories of language that characterize adult linguistic competence in much more child-friendly terms than does generative grammar—which makes the endpoint of language acquisition seem much closer.

The first point is that modern developmental psychologists and cognitive scientists no longer think of children's learning as isolated association-making and induction, but rather they think of it as integrated with other cognitive and social-cognitive skills—in ways that Skinner and the Behaviorists (and Chomsky in his critiques) could never have envisaged. Two sets of such skills are of particular importance for language acquisition. The first set comprises various skills of *intention-reading* (theory of mind, broadly conceived). These skills first emerge in human ontogeny at around 9–12 months of age (Tomasello, 1995a) and include such things as:

- the ability to share attention with other persons to objects and events of mutual interest (Bakeman and Adamson, 1984);
- the ability to follow the attention and gesturing of other persons to distal objects and events outside the immediate interaction (Corkum and Moore, 1995);
- the ability to actively direct the attention of others to distal objects by pointing, showing, and using of other nonlinguistic gestures (Bates, 1979);
- the ability to culturally (imitatively) learn the intentional actions of others, including their communicative acts underlain by communicative intentions (Tomasello, Kruger, and Ratner, 1993; Tomasello, 1998b).

These skills are necessary for children to acquire the appropriate use of any and all linguistic symbols, including complex linguistic expressions and constructions. Indeed, they basically define the symbolic or functional dimension of linguistic communication—which involves in all cases the attempt of one person to manipulate the intentional or mental states of other persons.* Importantly in the current context, this functional dimen-

* The notions of communicative intention and function are correlative. Someone uses a piece of language with a certain communicative intention, and so we may say that that piece of language has a certain function.

sion enables certain kinds of abstraction processes, such as analogy, that can only be effected when the elements to be compared play similar functional (communicative) roles in larger linguistic expressions and/or constructions. Intention-reading skills are very likely unique to human beings, and they probably emerged relatively recently in human evolution (Tomasello, 1999). They are domain-general in the sense that they do not just enable linguistic communication, but also enable a variety of other cultural skills and practices that children routinely acquire (such as tool use, pretend play, rituals).

The other main set of skills is those involved in various kinds of *pattern-finding*—categorization, broadly defined. These skills also begin to emerge early in human development (some prelinguistically) and include such things as:

- the ability to form perceptual and conceptual categories of "similar" objects and events (e.g., Rakison and Oakes, in press);
- the ability to form sensory-motor schemas from recurrent patterns of perception and action (e.g., Piaget, 1952; Schneider, 1999; Conway and Christiansen, 2001);
- the ability to perform statistically based distributional analyses on various kinds of perceptual and behavioral sequences (e.g., Saffran, Aslin, and Newport, 1996; Marcus et al., 1999; Gomez and Gerken, 1999; Ramus et al., 2000);
- the ability to create analogies (structure mappings) across two or more complex wholes, based on the similar functional roles of some elements in these different wholes (Gentner and Markman, 1997).

These skills are necessary for children to find patterns in the way adults use linguistic symbols across different utterances, and so to construct the grammatical (abstract) dimensions of human linguistic competence. They are skills that are evolutionarily fairly old, probably possessed in some form by all primates at the very least (Tomasello and Call, 1997; Hauser, Weiss, and Marcus, in press). They are also domain-general, in the sense that they allow organisms to categorize many different aspects of their worlds into a manageable number of kinds of things and events (although it seems very likely that when these skills are applied to linguistic symbols—as they are in humans but not in other primate species—some novel characteristics emerge). A particularly exciting development along these lines is the creation of connectionist and other kinds of computer programs that are able to find many patterns in linguistic stimuli with only a few uncomplicated pattern-finding algorithms (Elman, 1990, 1993). This of course suggests that young children should be able to do the same thing with similar skills—or even more with more skills.

The second modern development that undermines the You Can't Get

There From Here argument is new ways of looking at the nature of language itself. Chomskian generative grammar is a "formal" theory, meaning that it is based on the supposition that natural languages are like formal languages. Natural languages are thus characterized in terms of (1) a unified set of abstract algebraic rules that are both meaningless themselves and insensitive to the meanings of the elements they algorithmically combine, and (2) a lexicon containing meaningful linguistic elements that serve as variables in the rules. Principles governing the way the underlying algebra works constitute a universal grammar, the "core" of linguistic competence. The linguistic "periphery" involves such things as the lexicon, the conceptual system, irregular constructions and idioms, and pragmatics. This dichotomy between core and periphery leads to the so-called dual process approach to language acquisition (also called the words and rules approach by Pinker, 1999), namely, that whereas children acquire elements of the linguistic periphery using "normal" learning processes, the linguistic core, universal grammar, cannot be so learned; it is an innate property of the human mind.

But in recent years a new view of language and human linguistic competence has begun to emerge. This view is represented by a group of theories most often called *cognitive-functional linguistics* but sometimes also called *usage-based linguistics* to emphasize their central processing tenet that language structure emerges from language use (e.g., Langacker, 1987a, 1991, 2000; Croft, 1991, 2001; Goldberg, 1995; Givón, 1995; Bybee, 1985, 1995, 2002; see Tomasello, 1998a, in press, and Barlow and Kemmer, 2000, for similar approaches). Usage-based theories hold that the essence of language is its symbolic dimension, with grammar being derivative. The ability to communicate with conspecifics symbolically (conventionally, intersubjectively) is a species-specific biological adaptation. But, in contrast to generative grammar and other formal approaches, in usage-based approaches the grammatical dimension of language is a product of a set of historical and ontogenetic processes referred to collectively as *grammaticalization*. When human beings use symbols to communicate with one another, stringing them together into sequences, patterns of use emerge and become consolidated into grammatical constructions—for example, the English passive construction, noun phrase construction, or *-ed* past tense construction. As opposed to conceiving linguistic rules as algebraic procedures for combining words and morphemes that do not themselves contribute to meaning, this approach conceives linguistic constructions as themselves meaningful linguistic symbols—since they are nothing other than the patterns in which meaningful linguistic symbols are used in communication (for example, the passive construction is used to communicate about an entity to which something happens).

In the usage-based approach, competence with a natural language con-

sists of the mastery of all its items and structures, and these constitute a much more complex and diverse set of linguistic representations than the "core grammar" of formal approaches. They include the highly canonical (core), the highly idiosyncratic (periphery), and many things in between. Thus, fluent speakers of English control not only highly abstract syntactic constructions (past-tense *-ed,* the passive construction), but also concrete expressions based on individual words or phrases, such as ritualized greetings, idioms, metaphors, and noncanonical phrasal collocations (*I wouldn't put it past him; He's getting to me these days; Hang in there; That won't go down well with the boss; She put me up to it;* see Pawley and Syder, 1983; Jackendoff, 1996). In addition, and importantly, they also control many so-called mixed constructions that fall somewhere in between these in having both concrete and abstract elements (such as the *-er* construction, as in *The bigger they are, the nicer they are,* which has many unique properties along with some more regular ones). A plausible way to think of mature linguistic competence, then, is as a structured inventory of constructions, some of which are similar to many others and so reside in a more core-like center, and others of which connect to very few other constructions (and in different ways) and so reside more toward the periphery.

The implications of this new view of language for theories of language acquisition are truly revolutionary. If there is no clean break between the more rule-based and the more idiosyncratic items and structures of a language, then all constructions may be acquired with the same basic set of acquisitional processes—namely, those falling under the general headings of intention-reading and pattern-finding. If adult linguistic competence is based, to a much larger degree than previously supposed, on concrete pieces of language and straightforward generalizations across them—with many constructions remaining idiosyncratic and item-based into adulthood—then it is possible that children's early language is largely item-based and yet they can still construct an adult-like set of grammatical constructions originating with these baby constructions (given several years in which they hear several million adult utterances). If linguistic constructions are meaningful linguistic symbols in their own right, then children can use function or meaning to assist in their acquisition, just as they do in their acquisition of smaller linguistic constructions such as individual words.

In this book I proceed from the assumption that children can get from here to there, from item-based baby constructions to abstract constructions, and that they can do this with one set of acquisition processes. The assumption is justified by the fact that the cognitive and social learning skills that children bring to the acquisition process are much more power-

ful than previously believed, and by the fact that the adult endpoint of language acquisition comprises nothing other than a structured inventory of linguistic constructions, a much closer and more child-friendly target than previously believed. These two new advances in developmental psychology and usage-based linguistics thus encourage us to pursue the possibility that we might be able to describe and explain child language acquisition without recourse to any hypothesized universal grammar.

It should also be emphasized at the outset that, in the current view, the principles and structures whose existence it is difficult to explain without universal grammar (such Chomskian things as the subjacency constraint, the empty category principle, and the binding principles) are theory-internal affairs and simply do not exist in usage-based theories of language—full stop. There is no poverty of the stimulus when a structured inventory of constructions is the adult endpoint. Moreover, hypothesizing the existence of an innate universal grammar brings with it two major acquisition problems that are currently unresolved—and that do not exist on a usage-based view. First is the problem of cross-linguistic diversity: How can the child link her abstract universal grammar to the particularities of the particular language she is learning (the linking problem)? Second is the problem of developmental change: How can we understand the changing nature of children's language across development if universal grammar is always the same (the problem of continuity)? For these reasons as well, then, it seems worthwhile to attempt to describe and explain child language acquisition without adding the extra acquisitional problems created by an hypothesized universal grammar.

T W O

Origins of Language

The common behavior of mankind is the system of reference by means of which we interpret an unknown language.

—LUDWIG WITTGENSTEIN

HUMAN linguistic communication differs from the communication of other animal species in two main ways. First, and most importantly, human linguistic communication is symbolic. Linguistic symbols are social conventions by means of which one individual attempts to share attention with another individual by directing the other's attentional or mental state to something in the outside world. Other animal species do not communicate with one another using linguistic symbols, most likely because they do not understand that conspecifics have attentional or mental states that they could attempt to direct or share (Tomasello, 1998b). To oversimplify, animal signals are aimed at the behavior and motivational states of others, whereas human symbols are aimed at the attentional and mental states of others. It is this mental dimension that gives linguistic symbols their unparalleled communicative power, enabling them to be used to refer to and to predicate all kinds of diverse perspectives on objects, events, and situations in the world.

The second main difference is that human linguistic communication is grammatical. Human beings use their linguistic symbols together in patterned ways, and these patterns, known as linguistic constructions, take on meanings of their own—deriving partly from the meanings of the individual symbols but, over time, at least partly from the pattern itself. The process by which this occurs over historical time is called *grammaticalization* (or syntacticization), and grammatical constructions of course add still another dimension of communicative power to human languages. The process of grammaticalization depends crucially on a variety of domain-general cognitive and social-cognitive processes that operate as people communicate with and learn from one another. It is also a species-

unique process—because if other animals do not use symbols, the question of grammar is moot.

Human skills of linguistic communication are also unique in the way they are acquired during ontogeny. The main point is that, unlike other animal species, the human species does not have a single system of communication. Different groups of human beings have conventionalized different systems of communication (there are more than 6,000 of them), and children typically acquire only the system(s) of their natal group(s). Children take several years to acquire the many tens of thousands of linguistic conventions used by those around them, whereas most other animal species do not learn any of their species-typical communicative signals at all.

2.1. Phylogenetic Origins

As adumbrated in Chapter 1, the Generative Grammar hypothesis focuses only on grammar and claims that the human species has evolved during its phylogeny a genetically based universal grammar. The theory is unconcerned with the symbolic dimensions of human linguistic communication. The usage-based view—or at least the version of it espoused here—is precisely the opposite. In this view, the human use of symbols is primary, with the most likely evolutionary scenario being that the human species evolved skills enabling the use of linguistic symbols phylogenetically. But the emergence of grammar is a cultural-historical affair—probably originating quite recently in human evolution—involving no additional genetic events concerning language per se (except possibly some vocal-auditory information-processing skills that contribute indirectly to grammaticalization processes).

This is not to imply that we know how language originated in human evolution, because we do not. But if we focus on linguistic symbols as primary, we may obtain some hints by looking at the communication of our nearest primate relatives—who communicate not with symbols but with vocal and gestural signals. At the very least, this comparison will help us to identify the unique features of human symbolic communication. For hints about the emergence of grammar in human evolution we need to examine various processes of grammaticalization and syntacticization as they may be inferred from historical examinations of written language and from comparative examinations within language families.

2.1.1. Primate Communication

Discerning the unique features of human symbolic/linguistic communication is sometimes made more difficult by anthropocentric accounts of non-

human primate communication. The most important instance of this is the well-known case of the alarm calls of vervet monkeys. The basic facts are these (see Cheney and Seyfarth, 1990, for more details). In their natural habitats in east Africa vervet monkeys use three different types of alarm calls to indicate the presence of three different types of predator: leopards, eagles, and snakes. A loud, barking call is given to leopards and other cat species, a short cough-like call is given to two species of eagle, and a "chutter" call is given to a variety of dangerous snake species. Each call elicits a different escape response on the part of vervets who hear the call: to a leopard alarm they run for the trees; to an eagle alarm they look up in the air and sometimes run into the bushes; and to a snake alarm they look down at the ground, sometimes from a bipedal stance. These responses are just as distinct and frequent when researchers play back previously recorded alarm calls over a loudspeaker, indicating that the responses of the vervets are not dependent on seeing the predator but rather on information contained in the call itself.

On the surface, these alarm calls would seem to be very similar to human language. It seems as if the caller is directing the attention of others to something they do not perceive or something they do not know is present; that is, the calls would seem to be symbolic (referential). But several additional facts argue against this interpretation. First, there is basically no sign that vervet monkeys attempt to manipulate the attentional or mental states of conspecifics in any other domain of their lives. Thus, vervets also have different "grunts" that they use in various social situations, but these show no signs of being symbolic or referential in the sense of being intended to direct the attention of others to outside entities; they mainly serve to regulate dyadic social interactions not involving outside entities, such as grooming, playing, fighting, sex, and travel. Second, predator-specific alarm calls turn out to be fairly widespread in the animal kingdom. They are used by a number of species—from ground squirrels to domestic chickens—that must deal with multiple predators requiring different types of escape responses (Owings and Morton, 1998), but no one considers them to be symbolic or referential in a human-like way. An extremely important evolutionary fact in all of this is that no species of ape has such specific alarm calls or any other vocalizations that appear to be referential (Cheney and Wrangham, 1987). Since human beings are most closely related to the great apes, this means that it is not possible that vervet monkey alarm calls could be the direct precursor of human language unless at some point apes used them also—and there is no evidence of this.

Similarly and importantly, the visual-gestural communication of nonhuman primates shows no signs of referentiality or symbolicity either. Most strikingly, nonhuman primates do not point or gesture to outside objects

or events for others, they do not hold up objects to show them to others, and they do not even hold out objects to offer them to others (Tomasello and Call, 1997). Once again, primate gestures are used almost exclusively to regulate dyadic social interactions such as grooming, play, fighting, sex, and travel, not triadically to direct the attention of others to outside entities or events. Relatedly, nonhuman primates use their species-typical vocalizations and gestures almost exclusively for imperative motives, to request a behavior of others, not to share attention or information with others in a disinterested manner (Tomasello, 1998b).

Finally, nonhuman primate vocalizations and gestures are not socially learned in the sense of being copied from others. Primate vocalizations are almost certainly not learned at all, as monkeys and apes raised outside their normal social environments vocalize in much the same way as those who grow up in normal social environments (although some aspects of call comprehension and use may be learned). Many nonhuman primate gestures are also not learned, but some are. However, these are not learned by imitation—by observing others using a gesture and then adopting it oneself—but rather by a process of ritualization in which individuals mutually shape one another's behavior over repeated social interactions (Tomasello and Zuberbühler, 2002). Overall, because they are not used referentially, not used simply to share attention with others, and not learned from others via imitation, the communicative signals of nonhuman primates do not seem to be socially shared (or socially constituted) in the same way as human linguistic symbols.

As a result of facts such as these, a number of primatologists and behavioral ecologists have cautioned against using human language as an interpretive framework for nonhuman primate communication (Owings and Morton, 1998; Owren and Rendell, 2001). They concur with the current analysis that nonhuman primates do not use communicative signals to convey meaning or to convey information or to refer to things or to direct the attention of others, but rather use them to affect the behavior or motivational states of others directly. If this interpretation is correct, then the deep evolutionary roots of human language lie in the attempts of primate individuals to influence the behavior, not the mental states, of conspecifics. To find the most direct precursors of human linguistic symbols as tools for directing attention, therefore, we can only look at the history of the human species since it began its own unique evolutionary trajectory.

2.1.2. Symbols and Grammaticalization

Although no one knows for certain, it is very likely that human symbolic skills arose as a more or less direct result of a biological adaptation—most likely occurring very recently with the emergence of modern humans some

200,000 years ago. According to Deacon (1998), this adaptation concerned symbolic skills directly, whereas according to Tomasello (1999) it concerned a new kind of social cognition more generally, in which human beings understood one another for the first time as intentional and mental agents—which then led them to attempt to manipulate one another's intentional and mental states for various cooperative and competitive purposes.

In any case, whenever and however they arose, human linguistic symbols are most clearly distinguished from the communicative signals of other primate species by the ways they are learned and used:

- Human linguistic symbols are *socially learned,* mainly by cultural (imitative) learning in which the learner acquires not just the conventional form of the symbol but also its conventional use in acts of communication (Tomasello, Kruger, and Ratner, 1993).
- Because they are learned imitatively from others, linguistic symbols are understood by their users *intersubjectively* in the sense that users know their interlocutors share the convention (that is, everyone is potentially both a producer and a comprehender and they all know this; see Saussure, 1916, on "bi-directionality of the sign").
- Linguistic symbols are not used dyadically to regulate social interactions directly, but rather they are used in utterances *referentially* (triadically) to direct the attentional and mental states of others to outside entities (see Grice, 1975, on the non-natural meaning of linguistic symbols).
- Linguistic symbols are sometimes used *declaratively,* simply to inform other persons of something, with no expectation of an overt behavioral response (see Dunbar, 1996, on the origins of language for purposes of gossip).
- Linguistic symbols are fundamentally *perspectival* in the sense that a person may refer to one and the same entity as *dog, animal, pet,* or *pest,* or to the same event as *running, fleeing, moving,* or *surviving*—depending on her communicative goal with respect to the listener's attentional states (Langacker, 1987a).

All these features are in contrast to the unlearned, or at least not imitatively learned, dyadic and imperative communicative signals of nonhuman primates that do not involve mental perspectives at all. In at least one reasonable hypothesis, these uniquely human features all derive—along with a host of other cultural skills involving, for example, teaching and collaborative interactions—from a single social-cognitive adaptation enabling the understanding of the psychological states of others more generally (theory of mind, broadly defined; Tomasello, 1999).

Tomasello (1999) also argued that linguistic symbols provide human beings with a species-unique format for cognitive representation. That is, when a child learns the conventional use of linguistic symbols, what she is learning are the ways her forebears in the culture found it useful to share and manipulate the attention of others in the past. And because the people of a culture, as they move through historical time, evolve many and varied purposes for manipulating one another's attention (and because they need to do this in many different types of discourse situations), today's child is faced with a whole panoply of linguistic symbols and constructions that embody many different attentional construals of any given situation. As just a sampling, languages embody attentional construals based on such things as:

- Granularity-specificity *(thing, furniture, chair, desk chair).*
- Perspective *(chase-flee, buy-sell, come-go, borrow-lend).*
- Function *(father, lawyer, man, American; coast, shore, beach).*

Consequently, as the young child internalizes a linguistic symbol—as she learns the human perspective embodied in that symbol—she cognitively represents, not just the perceptual or motoric aspects of a situation, but also one way, among other ways of which she is also aware, that the current situation may be attentionally construed by "us," the users of the symbol. The way that human beings use linguistic symbols thus creates a clear break with straightforward perceptual or sensory-motor cognitive representations—even those connected with events displaced in space and/or time—and enables human beings to view the world in whatever way is convenient for the communicative purpose at hand.

The evolution of grammar raises a more controversial set of theoretical issues, leading to some very different hypothesized evolutionary scenarios. Generative grammarians believe that the human species evolved a genetically based universal grammar common to all peoples and that the variability in modern languages is basically on the surface only. There are a number of accounts from this perspective, ranging from Chomsky's (1986) single-mutation account to Bickerton's (1984) two-stage account to Pinker and Bloom's (1992) gradualist account. But in all these variants the basic idea is the same: that the fundamental grammatical categories and relations underlying all of the world's languages come from a biological adaptation (or set of adaptations) in the form of a universal grammar.

The alternative is the usage-based view, in which there is no need to posit a specific genetic adaptation for grammar because processes of grammaticalization and syntacticization can actually create grammatical structures out of concrete utterances—and grammaticalization and syntacticization are cultural-historical processes, not biological ones. Thus, it

is a historical fact that the specific items and constructions of a given language are not invented all at once, but rather they emerge, evolve, and accumulate modifications over historical time as human beings use them with one another and adapt them to changing communicative circumstances (Croft, 2000). Most importantly, through various discourse processes (involving various kinds of pragmatic inferencing, analogy making, and so on) loose and redundantly organized discourse structures congeal into more tightly and less redundantly organized constructions (see Traugott and Heine, 1991; Hopper and Traugott, 1993). This happens both on the level of words and on the level of more complex constructions.

On the level of words, simple examples are English phrases such as *on the top of* and *in the side of* evolving into *on top of* and *inside of* and eventually into *atop* and *inside*. Often, however, this congealing process results in some structural changes as the communicative functions of some elements are reanalyzed in the context of specific constructions. Thus, case markers and agreement markers most often originate in free-standing words such as spatial prepositions, pronouns, or even nouns and verbs. A simple English example concerns the future marker *gonna,* a fusion of *going* and *to*. The original use of *going* was as a verb for movement, often in combination with the preposition *to* to indicate the destination *(I'm going to the store),* but sometimes also to indicate an intended action that the *going to* enabled *(Why are you going to London? I'm going to see my bride)*. This later became *I'm gonna VERB,* with *gonna* indicating not just the intention to do something in the future, but futurity only (with no movement or intention necessary; on this change see Bybee, 2002). Givón's (1979) well-known characterization of this process is: today's morphology is yesterday's syntax.

On the level of constructions, instead of sequences of words becoming one word, whole phrases take on a new kind of organization; that is, loose discourse sequences become more tightly organized syntactic constructions. Again Givón's characterization is apt: today's syntax is yesterday's discourse. Some hypothetical examples based on Givón (although in many cases the historical record is not sufficiently detailed for confidence in the specifics):

- Loose discourse sequences such as *He pulled the door and it opened* may become syntacticized into *He pulled the door open* (a resultative construction).
- Loose discourse sequences such as *My boyfriend . . . He plays piano . . . He plays in a band* may become *My boyfriend plays piano in a band*. Or, similarly, *My boyfriend . . . He rides horses . . . He bets on*

them may become *My boyfriend, who rides horses, bets on them* (a relative clause construction).

- If someone expresses the belief that Mary will wed John, another person may respond with an assent, *I believe that,* followed by a repetition of the expressed belief, *Mary will wed John*—which become syntacticized into the single statement *I believe that Mary will wed John* (a sentential complement construction).
- Complex constructions may also derive from discourse sequences of initially separate utterances, as in *I want it . . . I buy it* evolving into *I want to buy it* (an infinitival complement construction).

The historical processes of grammaticalization and syntacticization derive from a number of psychological and social-communicative processes that have been well studied, most importantly automatization, functional reanalysis, and analogy. Thus, when a person says *going* and *to* together enough (and consistently for the same single function), she ends up saying *gonna* by processes of automaticity very similar to those which occur in a variety of sensory-motor skills (Schneider, 1999). The constraint on such streamlining is of course that the behavior cannot be so streamlined that it no longer serves its communicative function effectively. In situations of high predictability the reduction of phonetic content may be relatively great; in less predictable situations less reduction is possible without serious consequences for communication.

Frequency plays a large role in this process as well, as only relatively frequently used expressions will become highly predictable—which accounts for the well-known principle that the more frequently a word is used in a language the shorter it tends to be (Zipf's Law). Frequency is also crucial because, as is well known, constructions that occur frequently are often irregular. This irregularity can be maintained because items and constructions that are highly frequent can be learned and used on their own, as constructional islands, whereas items and constructions that are less frequent tend to get regularized by pattern-seeking children (or, in the limiting case, they drop out of use as children do not get enough exposures to learn them). An interesting example is the subjunctive in Canadian French, which has dropped out of active use for virtually all low-frequency verbs but has stayed in use for a small number of high-frequency verbs (Poplack, 2001; also note an even narrower pattern in English in which the subjunctive survives for most speakers only in some fixed expressions such as *If I were you . . .*).

Grammaticalization also quite often involves processes of functional reanalysis and analogy. An example from English illustrates (adapted from

Trask, 1996). Old English had a verb *lician* that meant something like "be pleasing to." Like similar verbs in many languages (such as the German *gefallen*, the Spanish *gustar*), this verb had as its subject the thing that pleased, with the person who was pleased with that item appearing in the dative case *(X is pleasing to Fred)*. The normal word order for utterances with this verb consisted of the person being pleased said before the verb (in dative case) and the thing doing the pleasing said after the verb (as subject, agreeing in number with the verb); this is presumably because in English nominals indicating people most often come before verbs (for pragmatic reasons of topicality) and nominals for inanimate objects most often come after verbs. We thus get:

Pam kynge licoden peran.
To the king-[dative] were-pleasing pears. (pears = plural subject)

During the Middle English period, however, English lost much of its case-marking morphology, and so this same utterance was normally expressed:

The king licenden peares.
The king were-pleasing pears. (no dative marking)

It is clear that *pears* is still the subject at this point since the verb agrees with it in number, and not with the singular *king* (the *-en* ending on the verb indicates plurality, as in modern-day German). Finally, the plural marking on the verb was lost too, and we were left with the modern-day:

The king liked pears.

The dative *king* has now been reanalyzed as the subject, and the former subject *pears* as a direct object. Presumably, a driving force in this particular historical development was the fact that this construction had an atypical configuration of case-marking and word order (and perhaps it became less frequent as well, creating pressure for regularization), and so the reanalysis was in some sense aided by a kind of analogy to other Subject-Verb-Object (SVO) constructions.

All of this is not perfectly understood at this point, but for the process of grammaticalization to result in complex and abstract syntactic constructions the organisms involved must be equipped with some fairly complex cognitive and social-cognitive skills, including the ability to form complex schemas and to categorize these and their internal constituents into abstract categories, as well as the abilities to make sophisticated pragmatic inferences, functional reanalyses, and analogies. It may also be that humans' relatively recent specialized speech adaptations enabled the emergence of fully linguistic communication, if for no other reason than that

they enabled the very rapid production of sequences of linguistic symbols so that grammaticalization could take place (Lieberman, 1985). In any case, grammaticalization theory is able, at least in principle, to account both for the similarities among the world's languages—based on species-wide skills of cognition, vocal-auditory information processing, and pragmatic inferencing, along with commonalities among peoples in social and communicative goals—and for fundamental differences in these languages, as different speech communities use and grammaticalize different discourse sequences.*

2.1.3. Language Universals

Of crucial importance to the question of whether human grammatical competence is best explained by an innate universal grammar or by processes of grammaticalization is the question of language universals. The basic facts are these. Leaving aside for the moment nouns and verbs—which may or may not be universal in all the world's languages—virtually all linguists who are involved in the detailed analysis of individual languages cross-linguistically (known as linguistic typologists) now agree that there are very few if any specific grammatical categories and constructions that are present in all languages. Many languages simply do not have one or more of what are conventionally called relative clauses, auxiliary verbs, passive constructions, grammatical markers for tense, grammatical markers of evidentiality, prepositions, topic markers, subject markers, a copula *(to be)*, case marking of grammatical roles, subjunctive mood, definite and indefinite articles, incorporated nouns, plural markers, conjunctions, adverbs, complementizers, and on and on. The fact is that many languages (or language families) have grammatical categories and constructions that seem to be unique to them, that is, that do not correspond to any of the European categories and constructions as these have been defined over the centuries, beginning with Greek and Roman sources—who, by the way, created these grammatical entities not with the goal of psychological real-

* Some people may doubt that cultural-historical processes can create abstract structures such as those embodied in the grammatical constructions of modern-day languages. But, although the analogy is clearly not perfect, there are many highly abstract structures in modern mathematics that could only have been created by cultural-historical processes since they are not universal among cultures (for example, those of algebra and calculus). Again, there are many disanalogies between language and mathematics (which is more closely related, both logically and historically, to written language). The only point is that abstract symbolic systems can be created by groups of human beings working together over historical time in the domain of mathematics, and so perhaps they can also be created in similar yet different ways in the domain of language.

about 12 months and not at six months or three years" (Bloom, 2000: 45). The puzzle is that infants seemingly have conceptualized things they can talk about from at least 4 or 5 months of age, by which time they have, by all accounts, formed concepts of simple objects and events (see, e.g., the research reviewed by Spelke et al., 1994; Baillargeon, 1995). Infants of about this same age have also demonstrated that they can recognize word-like sound patterns when these recur in their experience in association with distinct objects (Jusczyk, 1997). And, of course, one of the best-established findings in infancy research is that even neonates are able to associate two aspects of their experience with one another, including auditory and visual experiences (see Haith and Benson, 1997, for a review). But, since 5-month-olds do not comprehend or produce language, it would seem that concepts, speech units, and associations are not enough.

It is possible that further developments in infants' ability to conceptualize the world emerge at around the first birthday, and so account for the emergence of language. But this is unlikely. The only serious candidate in this regard is infants' emerging ability to deal with so-called sortal categories like "dog" and "duck" (e.g., Xu, Carey, and Welch, 1999). But such categories are not necessary for them to learn, for example, the proper names of those around them. It is also possible that further developments in infants' ability to segment speech are involved in the initial emergence of language. But, again, this is unlikely. Although infants are indeed acquiring new speech-perception skills at around their first birthdays, these would not be necessary for them to learn single words said to them in isolation, which occur with some frequency in at least some infants' daily lives well before language begins. And there exist no serious proposals that infants' skills of association learning undergo any kind of qualitative shift at 1 year of age that would provide some new boost to their ability to acquire language.

An alternative explanation involves infants' social and communicative skills. In this case something important does indeed seem to happen at the appropriate developmental period, and it does so in a way that is correlated with the emergence of language. Thus, although human infants are social creatures from very early in development—they look selectively at schematic drawings of human faces over other perceptual patterns (Fantz, 1963); they recognize other persons as animate beings that are different from physical objects (Legerstee, 1991); they engage in "protoconversations" with adults (Trevarthen, 1979); and they mimic some body movements (Meltzoff and Moore 1977, 1989, 1994)—they are probably not so different from other primate species socially. But near the end of the first year of life something new happens in the way human infants relate to

other persons, and, in the current account, this explains why the acquisition of language begins when it does.

In the current account, children begin to acquire language when they do because the learning process depends crucially on the more fundamental skills of joint attention, intention-reading, and cultural learning—which emerge near the end of the first year of life. And importantly, a number of studies have found that children's earliest skills of joint attentional engagement with their mothers correlate highly with their earliest skills of language comprehension and production (see Carpenter, Nagell, and Tomasello, 1998, for a review; and see Chapter 3 for studies of joint attention and word learning). This correlation derives from the simple fact that language is nothing more than another type—albeit a very special type—of joint attentional skill; people use language to influence and manipulate one another's attention.

2.2.2. Early Skills of Intention-Reading

At around 9–12 months of age human infants begin to engage in a host of new behaviors that would seem to indicate something of a revolution in the way they understand their social worlds. Prototypically, it is at this age that infants begin to flexibly and reliably look where adults are looking (gaze following), to use adults as social reference points (social referencing), and to act on objects in the way adults are acting on them (imitative learning). These behaviors are not dyadic—between child and adult (or child and object)—but rather they are triadic in the sense that they involve infants coordinating their interactions with both objects and people, resulting in a referential triangle of child, adult, and the object or event to which they share attention. These behaviors would seem to indicate an emerging understanding of other persons as intentional agents like the self whose psychological relations to outside entities may be followed into, directed, and shared (Tomasello, 1995a). Intentional agents are animate beings who have goals and who make active choices among behavioral means for attaining those goals, including active choices about what to pay attention to in pursuit of them.

Three manifestations of this new level of social understanding are especially important for language acquisition: (1) the joint attentional frame, (2) understanding communicative intentions, and (3) cultural learning in the form of role reversal imitation.

THE JOINT ATTENTIONAL FRAME

First, 1-year-olds' newfound ability to interact triadically with other persons enables them to participate in relatively extended bouts of social in-

teraction mediated by an object in which both participants constantly monitor each other's attention both to the object and to themselves. These periods of joint engagement establish the common ground—what we may call the joint attentional frame*—within which adult-child communication may take place. For example, suppose a child is on the floor playing with a toy, but also is perceiving many other things in the room. An adult enters the room and joins the child in playing with the toy. The joint attentional frame is those objects and activities that the child and the adult know are part of the attentional focus of both of them. In this case, such things as the rug and the sofa and the child's diaper will not be a part of the joint attentional frame, even though the child may be perceiving them basically continuously, because they are not part of "what we are doing." In contrast, if the adult enters the room with a new diaper and readies the child for a diaper change on the rug, then the joint attentional frame may include the diapers and perhaps the rug—but not the toys because "we" have no goals with respect to the toys.

The basic point is that joint attentional frames are defined intentionally, that is, they gain their identity and coherence from the child's and the adult's understandings of "what we are doing" in terms of the goal-directed activities in which we are engaged. In one case we are playing with a toy, which means that certain objects and activities are part of what we are doing, and in another case we are changing a diaper, which brings into existence, from the point of view of our joint attention, a whole different set of objects and activities. This enables the child, as we shall see shortly, to create the common ground within which she may understand the adult's communicative intentions when the adult uses a novel piece of language—at least partly by creating a domain of "current relevance." Another crucial feature of joint attentional frames is that the child understands both the adult's and her own roles in the interaction from the same "outside" perspective—so that they are all in a common representational format (Bruner, 1983; Tomasello, 1999).

UNDERSTANDING COMMUNICATIVE INTENTIONS

Second, 1-year-olds' newfound ability to understand others' communicative intentions enables them to understand communicative intentions inside these joint attentional frames. Human infants very likely begin to understand the intentional actions of others in the last few months of their first year of life, before language begins (Gergely et al., 1995). But commu-

* Other terms that have been used are "joint attentional formats" (Bruner, 1981) and "joint attentional scenes" (Tomasello, 1999).

nicative intentions are a special type of intention in which an individual intends something not just toward an inert object but toward the intention states of someone else. Consequently, when an adult addresses an utterance to an infant too young to comprehend intentions, from the infant's point of view the adult is just making noise (for whatever reason). Infants this young may on occasion learn to associate one of these noises with a perceptual event in much the same way a household pet may understand that the sound *dinner* heralds the arrival of food. But this is not language. Sounds become language for young children when and only when they understand that the adult is making that sound with the intention that they attend to something. This requires an understanding of other persons as intentional agents who intend things toward one's own intentional states.

To illustrate, Tomasello, Call, and Gluckman (1997) attempted to communicate with apes and human 2-year-olds by using communicative signs that were totally novel for the subjects. In two of their experimental conditions they indicated for subjects which of three distinct containers contained a reward by (a) placing a small wooden marker on top of the correct container, or (b) holding up an exact replica of the correct container. Before this experiment, children did not know about using markers and replicas as communicative signs, but they nevertheless used these novel signs very effectively to find the reward. In contrast, no ape was able to do this for either of the novel communicative signs. One explanation of these results is that the apes were not able to understand that the human being had intentions toward their attentional states. The apes therefore treated the communicative attempts of the human as discriminative cues on a par with all other types of discriminative cues that have to be laboriously learned over repeated experiences. The children, meanwhile, treated each communicative attempt as an expression of the adult's intention to direct their attention in ways relevant to their current situation.

Said another way, the children understood something of the experimenter's communicative intentions. In one reasonable analysis, to understand your communicative intention I must understand:

> You intend for [me to share attention to [X]]

Two aspects of this formulation are especially important. First, according to all analysts from Grice (1975) forward, the understanding of a communicative intention must have this embedded structure. Thus, if you physically push me down into a chair I will recognize your intention that I sit down. But if you tell me "Sit down" I will recognize your intention that I attend to your proposal that I sit down—and if I do sit down it will not be due to physical force but rather because I have changed my intentional

states to comply with your proposal. The understanding of a communicative intention is therefore a special case of the understanding of an intention; it is the understanding of another person's intention toward my intentional states. Understanding this is clearly more complex than understanding another person's intention *simpliciter.*

The other important aspect of this analysis is that it readily accommodates different kinds of speech act goals on the part of the speaker and their recognition by the listener. This is accomplished by simply substituting different things for the X in the formula. Thus, in the case of an imperative such as *Sit down,* I understand that you intend for me to attend to your proposal that I sit down. In the case of an indicative, referential utterance such as *A birdie!* I understand that you intend for me to share attention with you to the bird (to attend to your already established attention to the bird). Importantly, in the case of so-called performatives or expressives such as *Hi* or *Thank you,* I understand that you intend for me to attend to your expression of happiness at seeing me or your expression of gratitude at receiving this gift. The reason performatives are important in the current context is that most theories of language acquisition basically ignore them. But they are frequently used communicative symbols, and they have a very similar intentional structure to expressions with a more clearly referential component. If performatives were nothing more than spontaneous and unreflective expressions of emotion (with referential expressions involving some extra cognitive work), there would be no reason children could not begin using them at a much younger age than referential words—but they do not.

Children understand adult communicative intentions, including those expressed in linguistic utterances, most readily inside the common ground established by joint attentional frames. Using adults to highlight the general principles involved, suppose that an American is in a train station in Hungary when a native speaker approaches and starts talking to her in Hungarian. It is very unlikely that in this situation the American visitor will understand the communicative intentions expressed in any Hungarian word or phrase; there is no common ground or joint attentional frame. But suppose now that the American goes to the ticket window, manned by another Hungarian speaker, and tries to buy a ticket. In this situation it is possible that the visitor may come to comprehend the communicative intentions expressed in some Hungarian words and phrases because the two interactants share an understanding of each other's interactive goals in terms of gaining information about train schedules, obtaining a ticket, exchanging money, and so forth—goals expressed directly through the execution of meaningful and already understood actions such as the actual exchanging of ticket and money.

well, that is, they appeared to understand what the adult intended to do and performed that action instead of just mimicking the adult's actual behavior. In the second study, Carpenter, Akhtar, and Tomasello (1998a) investigated infants' imitation of accidental versus intentional actions. They had 16-month-olds watch an adult perform some two-action sequences on objects that made interesting results occur. One action of the modeled sequences was marked vocally as intentional ("There!"), and one action was marked vocally as accidental ("Woops!"). Infants were then given a chance to make the result occur themselves, and what they mainly did was to reproduce the adult's intentional actions but not the accidental ones. From soon after their first birthdays, then, infants cannot help perceiving Daddy as "trying to clean the table" or "trying to open the drawer"—not simply as making specific bodily motions or producing salient changes of state in the environment—and these intentional actions are what they attempt to reproduce.

Importantly, in learning to produce an act of symbolic communication, the process of imitative learning is similar to, but somewhat different from, the imitative learning of these straightforward intentional actions. For example, if the child sees an adult operate a novel toy in a particular way and then imitatively learns to do the same thing, there is a parallel in the way the adult and child treat the toy—the child just substitutes herself for the adult. However, when an adult addresses the child with a novel communicative symbol intending to refer her attention to that toy, and the child wants to imitatively learn this communicative behavior, the situation changes. The reason is that in expressing communicative intentions in a linguistic symbol, the adult expresses her intentions toward the child's attentional state. Consequently, if the child simply substitutes herself for the adult she will end up directing the symbol to herself—which is not what is needed. To learn to use a communicative symbol in a conventionally appropriate manner, the child must engage in role reversal imitation: she must learn to use a symbol toward the adult in the same way the adult used it toward her. This is clearly a process of imitative learning in which the child aligns herself with the adult in terms of both the goal and the means for attaining that goal; it is just that in this case the child must not only substitute herself for the adult as actor (which occurs in all types of cultural learning) but also substitute the adult for herself as the target of the intentional act (that is, she must substitute the adult's attentional state as goal for her own attentional state as goal).

The result of this process of role reversal imitation is a linguistic symbol: a communicative device understood intersubjectively from both sides of the interaction. That is to say, this learning process ensures that the child

understands that she has acquired a symbol that is socially "shared" in the sense that she can assume in most circumstances that the listener both comprehends and can produce that same symbol—and the listener also knows that they can both comprehend and produce the symbol (see Figure 2.2). This contrasts with the process of understanding communicative signals—for example, by nonhuman primates and presymbolic human infants—in which each participant understands its own role as sender or receiver only, from its own inside perspective. It is interesting to note that the intersubjectivity inherent in socially shared symbols, but not in one-way signals, sets up all kinds of pragmatic "implicatures" of the type investigated by Grice (1975) concerning expectations that other persons will use the conventional means of expression—that we both know they know—and not others that are more cumbersome or indirect.

The main thing to note in Figure 2.2 is the contrast between an associationistic account in which sounds are connected to objects (or concepts) in a direct, dyadic way and a social-pragmatic account in which the relationship is triadic and therefore not one of association but of intentionality (signifier-signified). Using linguistic symbols in utterances is a social act, and when this act is internalized in Vygotskian fashion the product is a unique kind of cognitive representation that is not only intersubjective (involving both self and other), but also perspectival in the sense that the child understands that the same referent could have been indicated in some other way—the speaker could have chosen another linguistic symbol to indicate a different aspect of this entity (Tomasello, 1999).

2.2.3. Early Skills of Pattern-Finding

In addition to these precursors for children's understanding of the symbolic dimensions of linguistic communication, prelinguistic infants demonstrate some of the prerequisite skills necessary for an understanding of the grammatical dimensions of linguistic communication. If we define these prerequisites as a pattern-finding skill (categorization, broadly defined), it has long been recognized that human infants are experts from early in development in finding visual patterns (see Haith and Benson, 1997, for a review). But some more recent findings have extended this to the auditory domain, and in some surprising ways.

It has recently been discovered that prelinguistic infants are able to find patterns in sequentially presented auditory stimuli with amazing facility. Saffran, Aslin, and Newport (1996) exposed 8-month-olds to two minutes of synthesized speech consisting of four tri-syllabic nonsense "words." For

nds that I share
ntion

Figure 2.2. Structure of a linguistic symbol. Each person can use it to intend (thick lines) that the partner follow her attention (thin lines) to some external entity, that is, to share attention to it.

example, infants would hear *bidakupadotigolabubidakutupiropadoti . . .* They were then exposed to two new streams of synthesized speech simultaneously (one presented to the left and one to the right) to see which they preferred to listen to (as indicated by the direction they turned their head). One of these streams contained "words" from the original (such as *tupiro* and *golabu*), whereas the other contained the same syllables but in a different order (so that there were no "words" from the original). Infants preferred to look toward the speech stream containing the "words" they had originally heard. The only cue in this experiment indicating "words" was that in the original and in one test stream the constituent syllables always occurred together (that is, the transition probabilities were equal to 1.0), whereas in the other test stream syllables occurred together randomly (that is, they never occurred together in the original; transition

probabilities were equal to 0). There were no other cues such as intonation or pauses or the like to indicate "word" boundaries.

Subsequent studies have shown that infants can find patterns even when the syllables from the original speech stream and the test speech stream are not the same. Marcus et al. (1999) found that 7-month-olds exposed repeatedly over a three-minute period to tri-syllabic nonsense "words" with the pattern ABB (such as *wididi, delili*) preferred in subsequent testing to look toward the speech stream containing other "words" having this same ABB pattern even though the specific syllables involved were totally new (such as *bapopo*). Gomez and Gerken (1999) found very similar results with 12-month-olds. These results indicate that prelinguistic infants are able to find patterns in auditory stimuli of an abstract nature, which would seem to be a necessary (although not sufficient) skill in the learning of abstract grammatical patterns in linguistic stimuli.

Two other sets of studies help to place these results into perspective. First, infants can find patterns of this same type in nonlinguistic tone sequences and even in visually presented sequences (Saffran et al., 1999; Kirkham, Slemmer, and Johnson, 2002). These pattern-finding skills are thus not specifically linguistic. Second, when nonhuman primates (specifically, tamarin monkeys) are tested in these same procedures, they show these same abilities (Ramus et al., 2000; Newport, Aslin, and Hauser, 2001; Hauser, Weiss, and Marcus, in press). These pattern-finding skills are thus not uniquely human, and so probably express very deep-seated skills of primate vocal-auditory processing. So it is important to remember that 7- and 8-month-old infants who are finding all of these patterns in auditory and visual stimuli in experiments do not process the grammatical constructions of real language—consisting of meaningful symbols—in either comprehension or production. Their pattern-finding skills are thus not sufficient by themselves for dealing with real grammatical constructions used for communication—because the infants do not comprehend the symbolic dimension of those constructions.

And so, what we have is an amazing set of necessary cognitive skills—namely, the statistical learning of concrete and abstract auditory patterns—that are ready to be put to use in constructing the grammatical dimensions of language, once children's ability to understand linguistic symbols comes on-line in the months surrounding their first birthdays. And interestingly, once language acquisition begins in earnest children use their pattern-finding skills on the functional (or meaning) side of things as well. That is, to learn the conventional use of a particular word the child not only must discern across instances that it is the same phonological form (the easiest, limiting case of pattern-finding) but also must see patterns in

the way adults use a particular form communicatively across different usage events. This functional pattern-finding ranges from seeing similarities in the different referents to which a word like *ball* might be applied to seeing similarities in the different relationships indicated by the many different uses of the word *for*.

2.3. Children's First Utterances

Intention-reading, broadly construed, is thus the foundational social-cognitive skill underlying children's comprehension of the symbolic dimensions of linguistic communication. Children begin to understand the linguistic symbols produced by adults when they are able to participate with adults in joint attentional frames and then, within that common ground, to understand their specific communicative intentions as expressed in an utterance. The ability to coordinate this intention-reading with social learning skills (creating cultural learning skills, including role reversal imitation) enables children to begin to acquire for themselves conventional linguistic symbols and a number of symbolically constituted gestures as well. With their skills in finding both concrete and abstract patterns in auditory sequences, once children have begun to acquire linguistic symbols they are also ready to begin relatively quickly to acquire more complex and abstract linguistic constructions. The motivational bases for all of this would seem to be specific to uniquely human social and cultural activities; in particular, the motivation would seem to emanate from (1) a desire to communicate with other persons, and (2) a desire to be like other persons (that is, to imitate them).

Children's first active uses of linguistic symbols take place within the common ground of joint attentional frames, and include both gestural and linguistic means. Most 1-year-olds produce a number of different kinds of gestures as well as some conventional linguistic symbols, and these two forms of communication are often coordinated in single utterances. Children of this age produce their gestural and linguistic utterances for both imperative motives, to get the adult to do something with respect to an object or event, and declarative motives, to get adults simply to share attention with them to some external event or entity (Bates, Camaioni, and Volterra, 1975). They also begin to make attempts to learn different kinds of symbols for expressing both aspects of their communicative intention that are already shared with their listener in the joint attentional frame (such as pronouns, demonstratives, some pointing) and aspects of their communicative intention that concern things outside that frame, which must be more specifically indicated (for example, with nouns and verbs).

tion or imitative learning or whether some infants learn in one way (especially prior to their first birthdays) and some learn in the other. And it may even happen that an infant who learns to point via ritualization later comes to comprehend adult pointing in a new way, and so comes to a new understanding of her own pointing and its equivalence to the adult version (Franco and Butterworth, 1996). Interestingly, Petitto (1988) has documented an important difference between the "natural" gestures of deaf children and their truly linguistic signs in American Sign Language (ASL). Most deaf children learn to point "naturally," but they also learn to point in ASL as symbols in this linguistic system (for example, for *me* and *you*). Deaf children differentiate these two types of pointing in several ways right from the beginning; for example, they sometimes make reversal errors with *me* and *you* as ASL symbols. These children thus seem to learn both an indexical or deictic form of pointing, as other children do, and also a symbolic form of pointing for ASL—most likely learned imitatively from observation of others using the ASL pointing symbol.

It is also of crucial theoretical significance that human infants point for others not just for imperative motives—to get help with something—but also for declarative motives such as simply wishing to share attention with them. Declarative pointing (and showing) may thus be the purest expression of the uniquely human social-cognitive motivation to share attention with others. Indeed, the lack of declarative pointing in the second year of life is a key diagnostic criterion for children with autism (Baron-Cohen, 1995).

The third kind of infant gestures is symbolic (sometimes called referential) gestures (Acredolo and Goodwyn, 1988; Pizzuto and Volterra, 2000). These are communicative acts that are associated with a referent either metonymically or iconically. Examples include such things as sniffing for a flower, panting for a dog, holding arms out for an airplane, raising arms for big things, and blowing for hot things. It is possible that some of these may be acquired via ritualization—the child performs a behavior spontaneously and the adult reacts in some positive way—but it is much more likely that in most cases infants are learning these symbolic gestures via imitation. That is, they are learning exactly as some infants learn to point symbolically via imitative learning or use linguistic symbols: by first understanding an adult's communicative intention in using the gesture and then engaging in role reversal imitation to use the gesture herself when she has "the same" communicative intention.

One interesting question concerning symbolic gestures is the role of iconicity. When the child holds out her arms like an airplane or pants like a dog, is she mimicking some aspect of the physical or behavioral proper-

ties of the object, or has she just learned from an adult a gesture that is as arbitrarily related to its referent as a linguistic symbol? There is not so much research relevant to this question, but it seems likely that the iconicity in such cases is in the eyes of the adult only and plays very little role in acquisition. Evidence for this interpretation is: (1) in the earliest stages, deaf children learning sign language are not helped by the iconicity of many sign language signs (Bonvillian, Garber, and Dell, 1997); (2) early in the second year, human infants can learn arbitrary gestures used referentially (like human object names) as easily as they learn words (Namy and Waxman, 1998); and (3) in experiments, 18-month-olds are unable to use iconicity to understand an adult's specific communicative intention (Tomasello, Striano, and Rochat, 1999). Symbolic gestures are thus very likely the same as spoken symbols in being learned via imitation of adults and in being only conventionally connected to their intended referents.

Thus, although human infants vocalize and babble from soon after birth, it is gestures that for many children seem to be the first carriers of their communicative intentions. And it is gestures that seem to pave the way to early language—at least from a functional point of view. In a study of the emergence of language in 12 Italian-speaking children, Iverson, Capirci, and Caselli (1994) found that virtually all the infants gestured frequently with adults, and that the function of children's gestures changed—from the primary carriers of communicative intent to a more supplementary function—as they began to acquire some conventional linguistic symbols (see also Marcos, 1991). Interestingly, in a comprehension experiment, Morford and Goldin-Meadow (1992) found that 1- and 2-year-old children could understand gestures in combination with speech, both when they were referentially redundant and when the gesture provided unique information (see also Golinkoff, 1983, on the interaction of speech and gesture in the early "negotiation of meaning" between infant and adult). Similarly, Harris, Barlow-Brown, and Chasin (1995) found a very strong correlation between children's tendency to point and their tendency to use object names. The research of Goodwyn and Acredolo (1993) also provides support for this position, as they found strong correlations between children's use of symbolic gestures before language and their early linguistic skills.

The importance and robustness of gesture as a communicative device are evidenced by the fact that even young blind children gesture while communicating (Iverson and Goldin-Meadow, 1998). And of course gesture remains a crucial aspect of human communication throughout childhood and even into adulthood (McNeill, 1992; Goldin-Meadow, 1997).

From this point of view, the existence of fully grammaticized sign languages and their ready acquisition by deaf children is not surprising. With respect to very early language in particular, it is interesting that deaf children acquiring a signed language do so on the same general timetable as hearing children learning a vocal language—thus demonstrating something of the robustness of the symbolic dimensions of human linguistic competence.

2.3.2. Early Holophrases

Most Western middle-class children begin producing conventional linguistic symbols in utterances in the months following their first birthdays. By the time they begin doing this, they typically have been communicating with other people gesturally and vocally for some months. Children's first linguistic expressions are learned and used in the context of these prior forms of nonlinguistic communication and for the same basic motives—declarative and imperative—and children soon learn to ask things interrogatively as well. There is typically a distinctive intonational pattern for each of these three types of speech act (declarative, imperative, interrogative). Children's first declarative utterances are sometimes about shared, topical referents and sometimes aimed at focusing the listener's attention on something new (typically assessed only from their own egocentric point of view; Greenfield and Smith, 1976).

At this early age the communicative functions of children's single-word utterances are an integral aspect of their reality for the child, and initially these functions (for example, imperative or interrogative) may not be well differentiated from the more referential aspects of the utterance (Ninio, 1992, 1993). That is to say, children's early one-word utterances may be thought of as "holophrases" that convey a holistic, undifferentiated communicative intention, most often the same communicative intention as that of the adult expressions from which they were learned (Barrett, 1982; Ninio, 1992). Many of children's early holophrases are relatively idiosyncratic, and their uses can change and evolve over time in a somewhat unstable manner. For example, Tomasello (1992a) reported the following holophrases for his daughter early in her language development:

- *Rockin:* First used while rocking in the rocking chair, then as a request to do so, and then as a name for the object.
- *Phone:* First used in response to hearing the telephone ring, then as she "talked" on the phone, then to point at and name the phone, and then when she wanted someone to pick her up so she could talk on the wall-phone (pointing to it).

- *Play-play:* First used as an accompaniment to her "playing" the piano, then to name the piano.
- *Towel:* First used as an accompaniment to her using a towel to clean up a spill, then to name the towel.
- *Steps:* First used as an accompaniment to her climbing or descending stairs (never to name the object).
- *Bath:* First used as an accompaniment to preparations for bath, then as she bathed her baby doll (never to name the object).
- *Game:* First used for others and then for herself playing with a baseball and baseball glove (never to name objects).
- *Make:* First used in block play to request that a structure be built, usually so that she could knock it down (and make a "mess").
- *Mess:* First used for the result of knocking down blocks, then when she wanted to knock them down.

In addition, however, some of children's holophrases are a bit more conventional and stable. Children speaking all the languages of the world often talk about such salient scenes of experience as the existence-nonexistence-recurrence of people and objects, the exchange-possession of objects, the movement-location of people and objects, various states and changes of states of objects, and the physical and mental activities of people (Brown, 1973). Thus, combining basic speech act motives and salient scenes of experience, young children of linguistic communities from around the world tend to use their earliest productive language to do such things as:

- request or indicate the existence of objects (for example, by naming them with a requestive or neutral intonation);
- request or describe the recurrence of objects or events (*more, again, another*);
- request or describe dynamic events involving objects (as described by *up, down, on, off, in, out, open, close*);
- request or describe the actions of people *(eat, kick, ride, draw);*
- comment on the location of objects and people *(here, outside);*
- ask some basic questions *(Whats-that? Where-go?)*;
- attribute a property to an object *(pretty, wet);* and
- use performatives to mark specific social events and situations *(hi, bye, thank you, no).*

An important issue for later language development is what parts of adult expressions children choose for their initial holophrases. The answer presumably lies in the specific language they are learning and the kinds of discourse in which they participate with adults, including the perceptual

salience of particular words and phrases in adults' speech (Slobin, 1985a). In English, most beginning language learners acquire a number of so-called relational words such as *more, gone, up, down, on,* and *off,* presumably because adults use these words in salient ways to talk about salient events (Bloom, Tinker, and Margulis, 1993; McCune, 1992). Many of these words are verb particles in adult English, and so the child at some point must learn to talk about the same events with phrasal verbs such as *pick up, get down, put on,* and *take off.* In Korean and Mandarin Chinese, in contrast, children learn fully adult verbs from the onset of language development because these verbs are most salient in adult speech to them (parallel to an English verb like *remove* for clothing; Choi and Gopnik, 1996; Gopnik and Choi, 1995; Tardif, 1996). When they begin with an adult verb as a holophrase, children must then at some point learn, at least for some discourse purposes, to fill in linguistically the nominal participants involved in the scene (as in *Remove shirt!*). Children in all languages also learn object labels for some events, such as *Bike!* as a request to ride a bicycle or *Birdie* as a comment on a passing flight, which means that they still need to learn to linguistically express the activity involved *(Ride bike! or See birdie).* The point is that children may begin talking about different scenes in different ways initially, and these ontogenetic starting points frame the subsequent task in particular ways.

In addition, most children begin language acquisition by learning some unparsed adult expressions as holophrases—such expressions as *I-wanna-do-it, Lemme-see,* and *Where-the-bottle.* The prevalence of this pattern in the early combinatorial speech of English-speaking children has been documented by Pine and Lieven (1993), who found that almost all children have at least some of these so-called frozen phrases in their early speech. This is especially true of some children, especially later-born children who observe siblings (Barton and Tomasello, 1994; Bates, Bretherton, and Snyder, 1988). In these cases there is different syntactic work to do if the child is to extract productive linguistic elements that can be used appropriately in other utterances, in other linguistic contexts, in the future. For this the child must engage in a process of segmentation, with regard not only to the speech stream but also to the communicative intentions involved—so as to determine which components of the speech stream go with which components of the underlying communicative intention.

As a nonlinguistic example, we may imagine that a child sees an adult use a stapler and understands that his goal is to staple together two pieces of paper. In some cases, the child may understand also that the sub-goal/function of placing the papers inside the stapler's jaws is to align them with the stapling mechanism inside the stapler, and that the sub-goal/function of pressing down on the stapler is to eject the staple through the

salient in the speech stream (they occur in utterance-final position, with stress, etc.). Other languages would seem to be more verb-friendly since many clauses consist of verbs only with no nominals (for example, when Chinese speakers indicate an ongoing event such as a boy kissing a girl, they quite often say only the equivalent of *Kiss*), and verbs are often more salient than nouns in the speech stream. Most critically, in basically all languages individual verbs—and many other relational words and function words—occur with higher token frequency in the language children hear than do nouns (since many relations and actions such as coming and going recur in the child's experience regularly, across many different situations, whereas particular objects such as ducks and flowers are mostly experienced irregularly). Nevertheless, children quite often, if not always, learn more nouns early in development than other types of words.

The claim that the so-called noun bias is universal has not gone unchallenged, however. A number of researchers have claimed that the hypothesis does not hold for particular languages, for example, Korean (Choi and Gopnik, 1995), Chinese (Tardif, 1996), and Tzotzill (de León, 2000). These are all very verb-friendly languages, and when spontaneous speech samples are taken the children quite often use more verbs than nouns early in development. The problem is that because children use each of their verbs more frequently than they use each of their nouns, spontaneous speech samples tend to underestimate children's noun vocabularies—since the probability that a child will use any particular noun in one hour of sampling is not very high. For this reason, Caselli, Casadio, and Bates (1999) used a parent interview measure to estimate the vocabularies of English-speaking children and Italian-speaking children, reasoning that this measure would be less sensitive to sampling issues. Italian has some of the properties of a verb-friendly language (e.g., verbs occur quite often at the ends of utterances in child directed speech) and so might be expected to show a verb advantage. But it did it not, and indeed Italian children show almost as strong a noun advantage as American children. Tardif, Gelman, and Xu (1999) addressed this issue directly by measuring Chinese children's vocabularies in both ways (spontaneous sample and parent interview), and the verb advantage for these children mostly disappeared with the interview measure. Surprisingly, there has been very little experimental work on this issue, but the studies that exist show that with similar numbers of exemplars children tend to learn novel nouns more easily than novel verbs (Goldin-Meadow, Seligman, and Gelman, 1976; Childers and Tomasello, in press).

Gentner (1982) provided a plausible explanation for the developmental priority of nouns: the Natural Partitions hypothesis. In brief, her hypothesis was that the nouns children learn early in development are proto-