

Constructing

a Language

A Usage-Based Theory
of Language Acquisition

MICHAEL TOMASELLO

HARVARD UNIVERSITY PRESS
Cambridge, Massachusetts, and London, England 2003

Copyright © 2003 by Michael Tomasello
All rights reserved
Printed in the United States of America

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


1. Language acquisition. 2. Cognition in children I. Title.

PT18 .T558 2003

401'.93—dc21 2002038840

Designed by Gwen Nefsky Frankfeldt

For Liz Bates



Constructing a Language

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

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 as Chomskian things as the subadjacency 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.

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