

Origins of Human Communication

Michael Tomasello

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

© 2008 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please e-mail special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Palatino by SNP Best-set Typesetter Ltd., Hong Kong, and was printed and bound in the United States.

Library of Congress Cataloging-in-Publication Data

Tomasello, Michael.

Origins of human communication / Michael Tomasello.

p. cm.—(Jean Nicod lectures)

Includes bibliographical references and index.

ISBN 978-0-262-20177-3 (hardcover : alk. paper)

1. Language and languages—Origin. 2. Animal communication. I. Title.

P116.T66 2008

401—dc22

2007049249

10 9 8 7 6 5 4 3 2 1

1

A Focus on Infrastructure

What we call meaning must be connected with the primitive language of gestures.

—Wittgenstein, *The Big Typescript*

Walk up to any animal in a zoo and try to communicate something simple. Tell a lion, or a tiger, or a bear to turn its body like “this,” showing it what to do by demonstrating with your hand or body and offering a delicious treat in return. Or simply point to where you would like it to stand or to where some hidden food is located. Or inform it that a fearsome predator is lurking behind a bush by both pointing to the location and pantomiming the predator’s actions. They don’t get it. And it is not just that they are not interested or motivated or intelligent in their own way, but the fact is that you simply cannot *tell* animals anything, even nonverbally, and expect them to understand.

Human beings, of course, find such gestures as pointing and pantomiming totally natural and transparent:

just look where I am pointing and you will *see* what I mean. Indeed, even prelinguistic infants use and understand the pointing gesture, and in many social situations in which vocal language is not possible or practical—for example, across a crowded room or in a noisy factory—humans naturally communicate by pointing and pantomiming. Tourists manage to survive and interact effectively in many situations in foreign cultures, in which no one shares their conventional language, precisely by relying on such naturally meaningful forms of gestural communication.

My central claim in these lectures is that to understand how humans communicate with one another using a language and how this competence might have arisen in evolution, we must first understand how humans communicate with one another using natural gestures. Indeed, my evolutionary hypothesis will be that the first uniquely human forms of communication were pointing and pantomiming. The social-cognitive and social-motivational infrastructure that enabled these new forms of communication then acted as a kind of psychological platform on which the various systems of conventional linguistic communication (all 6,000 of them) could be built. Pointing and pantomiming were thus the critical transition points in the evolution of human communication, already embodying most of the uniquely human forms of social cognition and motivation required for the later creation of conventional languages.

The problem is that, compared with conventional human languages (including conventionalized sign languages), natural gestures would seem to be very weak communicative devices, as they carry much less information “in” the communicative signal itself. Consider pointing, which I will argue later was the primordial form of uniquely human communication. Suppose that you and I are walking to the library, and out of the blue I point for you in the direction of some bicycles leaning against the library wall. Your reaction will very likely be “Huh?,” as you have no idea which aspect of the situation I am indicating or why I am doing so, since, by itself, pointing means nothing. But if some days earlier you broke up with your boyfriend in a particularly nasty way, and we both know this mutually, and one of the bicycles is his, which we also both know mutually, then the exact same pointing gesture in the exact same physical situation might mean something very complex like “Your boyfriend’s already at the library (so perhaps we should skip it).” On the other hand, if one of the bicycles is the one that we both know mutually was stolen from you recently, then the exact same pointing gesture will mean something completely different. Or perhaps we have been wondering together if the library is open at this late hour, and I am indicating the presence of many bicycles outside as a sign that it is.

It is easy to say that what carries the meaning in these different examples is “context,” but that is not very

helpful since all of the physical features of the immediate communicative context were (by stipulation) identical in the various scenarios. The only difference was our shared experience beforehand, and that was not the actual content of the communication but only its background. And so our question is: how can something as simple as a protruding finger communicate in such complex ways, and do so in such different ways on different occasions?

Any imaginable answer to this question will have to rely heavily upon cognitive skills of what is sometimes called mindreading, or intention-reading. Thus, to interpret a pointing gesture one must be able to determine: what is his intention in directing my attention in this way? But to make this determination with any confidence requires, in the prototypical instance, some kind of joint attention or shared experience between us (Wittgenstein's [1953] forms of life; Bruner's [1983] joint attentional formats; Clark's [1996] common conceptual ground). For example, if I am your friend from out of town and there is no way I could be familiar with your ex-boyfriend's bicycle, then you will not assume that I am indicating it for you. This is true even if, by some miracle, I do indeed know that this is his bicycle, but you do not know that I know this. In general, for smooth communication it is not enough that you and I each know separately and privately that this is his bicycle (and even that the other knows this); rather, this fact must be mutually known common ground between us. And in the case in which it is common ground between us that this is his bicycle, but

not that the two of you have just broken up (even if we each know this privately), then you will probably think that I am indicating your boyfriend's bicycle as a way of encouraging our entrance into the library, not discouraging it. The ability to create common conceptual ground—joint attention, shared experience, common cultural knowledge—is an absolutely critical dimension of all human communication, including linguistic communication with all of its *he's*, *she's*, and *it's*.

The other remarkable aspect of this mundane example of human pointing, from an evolutionary perspective, is its prosocial motivation. I am informing you of your ex-boyfriend's likely presence or the location of your stolen bicycle simply because I think you would want to know these things. Communicating information helpfully in this way is extremely rare in the animal kingdom, even in our closest primate relatives (in chapter 2 we will deal with examples such as warning cries and food calls). Thus, when a whimpering chimpanzee child is searching for her mother, it is almost certain that all of the other chimpanzees in the immediate area know this. But if some nearby female knows where the mother is, she will not tell the searching child, even though she is perfectly capable of extending her arm in a kind of pointing gesture. She will not tell the child because her communicative motives simply do not include informing others of things helpfully. In contrast, human communicative motives are so fundamentally cooperative that not only do we inform others of things helpfully, but one of the major ways we

request things from others is simply to make our desire known in the expectation that they will volunteer help. Thus, I may request a drink of water by simply stating that I want one (informing you of my desire), knowing that, in most instances, your tendency to be helpful (and our mutual knowledge of this) turns this act of informing into what is effectively a full-blown request.

Human communication is thus a fundamentally cooperative enterprise, operating most naturally and smoothly within the context of (1) mutually assumed common conceptual ground, and (2) mutually assumed cooperative communicative motives. The fundamentally cooperative nature of human communication is, of course, the basic insight of Grice (1957, 1975), and it is assumed—to varying degrees and in various ways—by others who follow in this tradition such as Clark (1992, 1996), Sperber and Wilson (1986), and Levinson (1995, 2006). But if we are to understand the ultimate origins of human communication, both phylogenetically and ontogenetically, we must look outside of communication itself and into human cooperation more generally. It turns out that human cooperation is unique in the animal kingdom in many ways, both structurally and motivationally.

Specifically, human cooperation is structured by what some modern philosophers of action call shared intentionality or “we” intentionality (Searle 1995; Bratman 1992; Gilbert 1989). In general, shared intentionality is what is necessary for engaging in uniquely human forms of collaborative activity in which a plural subject “we” is

involved: joint goals, joint intentions, mutual knowledge, shared beliefs—all in the context of various cooperative motives. The jointness involved is especially salient in institutional interactions involving such culturally constructed entities as money, marriage, and government, which exist only within an institutional reality, collectively constituted, in which we all believe and act together as if they do exist. But shared intentionality is involved in simpler and more concrete collaborative activities as well, for example, when we form the shared goal to construct a tool together or to take a walk together, or when we simply admire a mountain vista together or engage in a religious practice together. The proposal is thus that human cooperative communication—whether using “natural” gestures or “arbitrary” conventions—is one instance, albeit a special instance, of uniquely human cooperative activity relying on shared intentionality (Tomasello, Carpenter, Call, Behne, and Moll 2005). The skills and motivations of shared intentionality thus constitute what we may call the cooperative infrastructure of human communication.

If human communication is cooperatively structured in ways that the communication of other primates is not, the question naturally arises how it could have evolved. The issue is that in modern evolutionary theory the emergence of cooperation, or at least altruism, is always problematic. But if the infrastructure of human cooperative communication is basically the same as that of all other collaborative activities, then one possibility is that it

evolved as part of a larger human adaptation for cooperation and cultural life in general. Thus, for reasons we do not know, at some point in human evolution individuals who could engage with one another collaboratively with joint intentions, joint attention, and cooperative motives were at an adaptive advantage. Cooperative communication then arose as a way of coordinating these collaborative activities more efficiently, first inheriting and then helping to build further a common psychological infrastructure of shared intentionality. This all began almost certainly in mutualistic activities in which an individual who helped her partner was simultaneously helping herself. But then there was a generalization to more altruistic situations in which individuals simply informed or shared things with others freely, possibly as a way to cultivate reciprocity and a reputation for cooperation within the cultural group. Only later still did humans begin to communicate in this new cooperative way outside of cooperative contexts for higher-up, noncooperative purposes—leading to the possibility of deception by lying.

The initial steps in this process almost certainly took place in the gestural modality. This becomes especially clear when we compare the vocal and gestural communication of our nearest primate relatives, the great apes. Great ape vocalizations are almost totally genetically fixed, based on almost no learning, tightly tied to specific emotions, and broadcast indiscriminately to everyone in the immediate vicinity. In contrast, many great ape ges-

tures are learned and used quite flexibly in different social circumstances for different social ends—with new gestures sometimes learned for interacting with humans—and communicators direct these gestures at specific individuals taking into account their current attentional state. Learning, flexibility, and attention to the partner are obviously fundamental characteristics of the human way of communicating, and things simply could not move in the human direction until they were present. It is also important, as many gestural origins theorists have noted previously, that the human use of pointing and pantomiming—as the successors to ape gestures after things became cooperative—are “natural” in a way that “arbitrary” linguistic conventions are not. Specifically, pointing is based on humans’ natural tendency to follow the gaze direction of others to external targets, and pantomiming is based on humans’ natural tendency to interpret the actions of others intentionally. This naturalness makes these gestures good candidates as an intermediate step between ape communication and arbitrary linguistic conventions.

And what about language? The current hypothesis is that it is only within the context of collaborative activities in which participants share intentions and attention, coordinated by natural forms of gestural communication, that arbitrary linguistic conventions could have come into existence evolutionarily. Conventional languages (first signed and then vocal) thus arose by piggybacking on these already understood gestures, substituting for the

naturalness of pointing and pantomiming a shared (and mutually known to be shared) social learning history. This process was, of course, made possible by humans' unique skills of cultural learning and imitation, which enable them to learn from others and their intentional states in uniquely powerful ways (Tomasello 1999). As part of this same evolutionary trajectory, human beings also began to create and pass along culturally various grammatical conventions organized into complex linguistic constructions that codified complex *types* of messages for use in recurrent communicative situations.

We thus need basic evolutionary processes, working in several different ways, to explain the origin of the underlying psychological infrastructure of human cooperative communication. But then in addition, to explain the origins of humans' 6,000 different conventional languages, we also need cultural-historical processes in which particular linguistic forms are conventionalized in particular speech communities, and then sequences of these are grammaticalized into grammatical constructions, and then all of these conventions and constructions are passed along to new generations via cultural learning. We thus may see here especially clearly the ongoing dialectic between evolutionary and cultural-historical processes as first described by Vygotsky (1978) and, in a more modern evolutionary framework, by Richerson and Boyd (2005)—and with which I myself have been obsessed for some time (Tomasello, Kruger, and Ratner 1993; Tomasello 1999; Tomasello et al. 2005). This perspective on

human communication and language thus basically turns the Chomskian proposal on its head, as the most fundamental aspects of human communication are seen as biological adaptations for cooperation and social interaction in general, whereas the more purely linguistic, including grammatical, dimensions of language are culturally constructed and passed along by individual linguistic communities.

In all, the road to modern human communication was almost certainly a long and circuitous one, with many twists and turns along the way. To provide a theoretical account based mainly on empirical data, then, we must consider many different aspects of ape and human life—which makes this account a long and circuitous one as well. But despite the many complexities along the way, our final destination is easily stated and crystal clear: identification of the species-unique features of human communication and their ontogenetic and phylogenetic roots. Toward this end, in what follows I evaluate three specific hypotheses:

1. Human cooperative communication emerged first in evolution (and emerges first in ontogeny) in the natural, spontaneous gestures of pointing and pantomiming.
2. Human cooperative communication rests crucially on a psychological infrastructure of shared intentionality, which originated evolutionarily in support of collaborative activities, and which comprises most importantly:
 - (a) social-cognitive skills for creating with others

joint intentions and joint attention (and other forms of common conceptual ground), and (b) prosocial motivations (and even norms) for helping and sharing with others.

3. Conventional communication, as embodied in one or another human language, is possible only when participants already possess:

(a) natural gestures and their shared intentionality infrastructure, and

(b) skills of cultural learning and imitation for creating and passing along jointly understood communicative conventions and constructions.