



Michael Tomasello

Die Ursprünge

der menschlichen

Kommunikation

Suhrkamp

SV

Michael Tomasello
Die Ursprünge der
menschlichen
Kommunikation

Aus dem Amerikanischen von
Jürgen Schröder

Suhrkamp

Titel der Originalausgabe: *Origins of Human Communication*.

Erstmals veröffentlicht 2008.

Die Übersetzung erscheint mit freundlicher Genehmigung von
MIT Press, Cambridge (Mass.)/London.

© 2008 Massachusetts Institute of Technology

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Erste Auflage 2009

© der deutschen Ausgabe Suhrkamp Verlag Frankfurt am Main
2009. Alle Rechte vorbehalten, insbesondere das des öffentlichen
Vortrags sowie der Übertragung durch Rundfunk und Fernsehen,
auch einzelner Teile. Kein Teil des Werkes darf in irgendeiner
Form (durch Fotografie, Mikrofilm oder andere Verfahren)
ohne schriftliche Genehmigung des Verlages reproduziert oder
unter Verwendung elektronischer Systeme verarbeitet, verviel-
fältigt oder verbreitet werden.

Druck: Druckhaus Nomos, Sinzheim

Printed in Germany

ISBN 978-3-518-58538-2

3 4 5 6 7 8 - 15 14 13 12 11 10

Zeig auf ein Stück Papier! – Und nun zeig auf
seine Form, – nun auf seine Farbe, nun auf seine
Anzahl ... – Nun, wie hast du es gemacht?

Ludwig Wittgenstein,

Philosophische Untersuchungen

Inhalt

Vorwort und Danksagung	9
1 Konzentration auf die Infrastruktur	12
2 Intentionale Kommunikation bei Primaten	24
2.1 Stimmliche Displays	26
2.2 Gestische Signale	31
2.3 Kommunikation mit Menschen	46
2.4 Intentionalität bei der Kommunikation von Affen	56
2.5 Schlußfolgerung	66
3 Kooperative Kommunikation beim Menschen	68
3.1 Zeigegesten und Gebärdenspiel	71
3.2 Das Kooperationsmodell	83
3.3 Kommunikationskonventionen	112
3.4 Schlußfolgerung	119
4 Ontogenetische Ursprünge	121
4.1 Zeigegesten bei Kleinkindern	123
4.2 Quellen der Zeigegesten von Kleinkindern	148
4.3 Frühes Gebärdenspiel	159
4.4 Geteilte Intentionalität und frühe Sprache	168
4.5 Schlußfolgerung	180
5 Phylogenetische Ursprünge	183
5.1 Die Entstehung von Zusammenarbeit	186
5.2 Das Entstehen kooperativer Kommunikation	206
5.3 Die Entstehung konventioneller Kommunikation	234
5.4 Schlußfolgerung	254

6 Die grammatische Dimension	260
6.1 Die Grammatik des Aufforderns	263
6.2 Die Grammatik des Informierens	288
6.3 Die Grammatik des Teilens und der Erzählung ...	301
6.4 Die Konventionalisierung von Sprach- konstruktionen	316
6.5 Schlußfolgerung	336
7 Von Affengesten zur Sprache des Menschen	339
7.1 Zusammenfassung des Arguments	339
7.2 Hypothesen und Probleme	347
7.3 Sprache als geteilte Intentionalität	362
Literatur	367
Namenregister	390
Sachregister	395

Vorwort und Danksagung

Dieses Buch beruht auf den Jean-Nicod-Vorlesungen, die ich im Frühjahr 2006 in Paris gehalten habe. Um den Interessen der Wissenschaftler, die am Institut Jean Nicod arbeiten, entgegenzukommen, entschloß ich mich dazu, das Thema »Kommunikation« in den Mittelpunkt zu stellen. Ich habe umfangreiche theoretische und empirische Arbeiten durchgeführt, erstens zur gestischen Kommunikation bei Menschenaffen; zweitens zur gestischen Kommunikation bei Kleinkindern; und drittens zur frühen Sprachentwicklung bei Kindern. Außerdem habe ich ausführlich über allgemeinere kognitive und sozio-kognitive Prozesse geforscht, die an menschlicher Kommunikation und Sprache beteiligt sind: (1) soziale und kulturelle Kognition; (2) soziales und kulturelles Lernen; und (3) Kooperation und geteilte Intentionalität. In diesem Band unternehme ich den Versuch, all diese Dinge in einer zusammenhängenden Darstellung der Evolution und der Entwicklung menschlicher Kommunikation zusammenzuführen. Die leitende Idee, von der dieser Versuch inspiriert ist, besagt, daß es einige recht spezifische Verbindungen geben muß zwischen der grundlegend kooperativen Struktur menschlicher Kommunikation, wie sie ursprünglich von Grice entdeckt wurde, und der besonderen kooperativen Struktur sozialer Interaktion und Kultur im allgemeinen, wie sie für Menschen im Gegensatz zu anderen Primaten kennzeichnend ist.

Die Gedanken in diesem Buch stammen hauptsächlich aus der gemeinsamen Forschung und den Diskussionen mit meinen vielen Kollegen in der Abteilung für Vergleichende und Entwicklungspsychologie am Max-Planck-Institut für Evolutionäre Anthropologie. Vieles von dem, was hier vorgestellt wird, hat seinen Ursprung in diesem Forschungszusammen-

hang, und ich wünschte nur, ich könnte mich an all die einzelnen Quellen deutlicher erinnern.

Deutlich ist jedoch die große Schuld, in der ich bei verschiedenen Einzelpersonen stehe. An erster Stelle ist hier Malinda Carpenter zu nennen. Malinda und ich haben fast täglich über Themen, die sich mehr oder weniger direkt auf den vorliegenden Band beziehen, gesprochen. Durch diese Diskussionen wurde mein Denken so grundlegend geformt, daß es leider unmöglich ist, spezifische Anerkennung für einzelne Gedanken zu zollen (oder alle diejenigen Punkte anzugeben, in denen Malinda mir nicht zustimmt). Von besonderer Bedeutung waren außerdem die vielen Gespräche, die ich über die Jahre hinweg mit Josep Call über die gestische Kommunikation von Menschenaffen und mit Elena Lieven über den Spracherwerb von Kindern geführt habe.

Eine frühere Fassung der in diesem Band enthaltenen Gedanken stellte ich den Mitgliedern unserer Forschungsgruppe zur sozialen Kognition (in den berüchtigten September-Sitzungen) vor. Dabei erhielt ich äußerst nützliche Rückmeldungen von Hannes Rakoczy, Tanya Behne, Henrike Moll, Ulf Liszkowski, Felix Warneken, Emily Wyman, Suse Grassmann, Kristin Liebal, Maria Gräfenhain, Gerlind Hauser und anderen. Darunter war auch der Vorschlag, eine Reihe von Diagrammen wegzulassen, die noch verwegener waren als diejenigen, die nun hier enthalten sind. Die Zuhörer der Jean-Nicod-Vorlesungen selbst machten ebenfalls eine Reihe nützlicher Vorschläge, insbesondere Dan Sperber.

Verschiedene Personen haben mehr oder weniger den gesamten Band gelesen und halfen mir dabei, ihn gewaltig zu verbessern: Malinda Carpenter, Elena Lieven, Bill Croft, Adele Goldberg und Gina Conti-Ramsden – zusammen mit einem anonymen Gutachter für MIT-Press. Andere haben ausgewählte Teile gelesen und gaben mir ebenfalls wertvolle Rückmeldungen. Dazu gehören Hannes Rakoczy, Henrike Moll,

Joe Henrich, Danielle Matthews, Nausicaa Pouscoulous, Felix Warneken, Colin Bannard, Emily Wyman und Kristin Liebal. Aufgrund der durchdachten Kritik dieser Leser wurde dieser Band empirisch exakter, theoretisch kohärenter und zudem benutzerfreundlicher. Ich danke auch Esteban Rivas für hilfreiche Informationen, die in *Tabelle 6.1* eingegangen sind.

Schließlich möchte ich wie immer Henriette Zeidler danken, die nicht nur bei verschiedenen Aspekten des Buchs eine Hilfe war, sondern auch die Angelegenheiten in der Abteilung gewohnt souverän und herzlich regelte, während ich zu Hause war und schrieb. Ich danke auch Annett Witzmann für ihre Hilfe bei den Literaturangaben und Tom Stone von MIT-Press für seine Betreuung des gesamten Veröffentlichungsprozesses.

1 Konzentration auf die Infrastruktur

Was wir Bedeutung nennen, muß mit der primitiven
Gebärdensprache (Zeigesprache) zusammenhängen.

Ludwig Wittgenstein, *The Big Typescript*

Versuchen Sie einmal, irgendeinem Tier im Zoo etwas Einfaches mitzuteilen. Teilen Sie einem Löwen, einem Tiger oder einem Bären mit, seinen Körper »so« zu drehen, indem Sie ihm durch eine Geste Ihrer Hand oder Ihres Körpers zeigen, was er tun soll, und ihm dafür eine köstliche Leckerei anbieten. Oder zeigen Sie einfach auf die Stelle, wo sich das Tier hinstellen soll oder das Futter versteckt ist. Oder informieren Sie es darüber, daß ein furchterregendes Raubtier hinter einem Busch lauert, indem Sie sowohl auf die Stelle zeigen als auch durch Gebärdenspiel die Handlungen des Raubtiers nachahmen. Das Tier wird Sie nicht verstehen. Das liegt nicht daran, daß es nicht interessiert oder motiviert oder auf seine eigene Art intelligent ist, sondern daran, daß Sie Tieren einfach nichts *mitteilen* können, nicht einmal nonverbal, und auch nicht erwarten können, daß sie das Gesagte verstehen.

Menschen finden solche Gesten wie Zeigen und Gebärdenspiel gewiß völlig natürlich und durchsichtig: Schauen Sie einfach, wohin ich zeige, und Sie werden *sehen*, was ich meine. Tatsächlich verwenden und verstehen sogar Kleinkinder, die noch nicht sprechen können, die Zeigegeste, und in vielen sozialen Situationen, in denen verbale Verständigung unmöglich oder unpraktisch ist – beispielsweise über einen überfüllten Raum hinweg oder in einer lauten Fabrik –, kommunizieren Menschen natürlicherweise durch Zeigegesten und Gebärdenspiel. Touristen schaffen es in vielen Situationen innerhalb fremder Kulturen, in denen niemand ihre konventionelle Sprache teilt, zu überleben und erfolgreich zu

interagieren, indem sie sich gerade auf solche von Natur aus bedeutungstragenden Formen gestischer Kommunikation stützen.

Meine zentrale Behauptung in diesem Buch ist, daß wir zuerst verstehen müssen, wie Menschen durch den Gebrauch natürlicher Gesten miteinander kommunizieren, bevor wir nachvollziehen können, wie Menschen durch den Gebrauch einer Sprache miteinander kommunizieren und wie diese Fertigkeit im Lauf der Evolution entstanden sein könnte. Meine evolutionäre Hypothese wird nämlich lauten, daß die ersten, nur beim Menschen vorkommenden Formen der Kommunikation im Zeigen und Gebärdenspiel bestanden. Die sozio-kognitive und sozio-motivationale Infrastruktur, die diese neuen Formen der Kommunikation ermöglichte, wirkte dann als eine Art psychologische Plattform, auf der die verschiedenen Systeme konventioneller sprachlicher Kommunikation (alle 6000 existierenden) aufgebaut werden konnten. Zeigen und Gebärden waren somit die entscheidenden Übergangspunkte in der Evolution menschlicher Kommunikation und beinhalteten schon die meisten der nur beim Menschen vorkommenden Formen sozialer Kognition und Motivation, die für die spätere Schaffung konventioneller Sprachen erforderlich waren.

Das Problem ist, daß natürliche Gesten im Vergleich zu konventionellen menschlichen Sprachen (einschließlich konventionalisierter Zeichensprachen) sehr schwache Kommunikationsmittel zu sein scheinen, da sie viel weniger Information »im« Kommunikationssignal selbst enthalten. Betrachten wir das Zeigen, das, wie ich später argumentieren werde, die ursprüngliche Form der nur beim Menschen vorkommenden Kommunikation war. Nehmen wir an, Sie und ich seien auf dem Weg zur Bibliothek, und aus heiterem Himmel zeige ich in die Richtung einiger Fahrräder, die an der Mauer der Bibliothek lehnen. Ihre Reaktion wird sehr wahrscheinlich

»Häh?« sein, weil Sie keine Ahnung davon haben, auf welchen Aspekt der Situation ich hinweise oder warum ich das tue, da das Zeigen an sich nichts bedeutet. Aber wenn Sie sich einige Tage zuvor auf besonders üble Weise von Ihrem Freund getrennt haben, und wir beide wissen, daß der andere das weiß, und eines der Fahrräder ihm gehört, wovon wir ebenfalls wissen, daß der andere das weiß, dann könnte genau dieselbe Zeigegeste in genau derselben Situation etwas sehr Komplexes bedeuten wie zum Beispiel »Ihr Exfreund ist schon in der Bibliothek (wir sollten sie deshalb vielleicht meiden)«. Wenn andererseits eines der Fahrräder dasjenige ist, von dem wir beide wissen, daß der andere weiß, daß es Ihnen vor kurzem gestohlen wurde, dann wird genau dieselbe Zeigegeste etwas völlig anderes bedeuten. Oder vielleicht haben wir uns gefragt, ob die Bibliothek zu dieser späten Stunde noch geöffnet ist, und ich weise auf die vielen Fahrräder draußen hin als Zeichen dafür, daß sie geöffnet ist.

Man könnte sagen, daß das, was die Bedeutung in diesen verschiedenen Beispielen trägt, der »Kontext« ist, aber das ist nicht sehr hilfreich, da alle physischen Merkmale des unmittelbaren kommunikativen Kontextes in den verschiedenen Szenarien (per Festsetzung) identisch sind. Unterschiedlich waren jeweils und einzig unsere vorgängigen gemeinsamen Erfahrungen, und diese waren nicht der eigentliche Inhalt der Kommunikation, sondern nur ihr Hintergrund. Unsere Frage lautet daher: Wie kann etwas so Einfaches wie ein ausgestreckter Finger auf so komplexe Weise etwas mitteilen – und das bei verschiedenen Gelegenheiten auf so verschiedene Weise?

Jede denkbare Antwort auf diese Frage wird sich stark auf kognitive Fertigkeiten beziehen müssen, die manchmal Fertigkeiten des Erfassens geistiger Zustände oder des Erfassens von Intentionen genannt werden. Um eine Zeigegeste zu interpretieren, muß man also feststellen können, welche

Absicht der andere mit einer derartigen Lenkung meiner Aufmerksamkeit verfolgt. Um aber diese Feststellung einigermaßen zuverlässig zu treffen, ist im prototypischen Fall eine Art von gemeinsamer Aufmerksamkeit oder geteilter Erfahrung erforderlich (Wittgensteins Lebensformen; Bruners Formate gemeinsamer Aufmerksamkeit; Clarks gemeinsamer begrifflicher Hintergrund).¹ Wenn ich beispielsweise einer Ihrer Freunde bin, nicht in derselben Stadt lebe und es unmöglich ist, daß ich mit dem Fahrrad Ihres Exfreundes vertraut bin, dann werden Sie nicht annehmen, daß ich Sie darauf hinweise. Das ist selbst dann so, wenn ich wundersamerweise tatsächlich wissen sollte, daß dies sein Fahrrad ist, Sie aber nicht wissen, daß ich das weiß. Im allgemeinen genügt es für eine reibungslose Kommunikation nicht, daß Sie und ich gesondert und privat wissen, daß dies sein Fahrrad ist (es genügt nicht einmal, daß der andere das weiß). Diese Tatsache muß vielmehr ein wechselseitig zwischen uns geteiltes Wissen sein. Und wenn wir ein geteiltes Wissen darüber besitzen, daß dies sein Fahrrad ist, aber nicht, daß Sie und Ihr Freund sich gerade getrennt haben (selbst wenn jeder von uns das privat weiß), dann werden Sie wahrscheinlich denken, daß ich auf das Fahrrad Ihres Exfreundes hinweise, um zum Betreten der Bibliothek anzuregen, und nicht, um Sie davon abzubringen. Die Fähigkeit, einen gemeinsamen begrifflichen Hintergrund² zu schaffen – gemeinsame Aufmerksamkeit, geteilte Erfahrung, gemeinsames kulturelles Wissen –,

1 Wittgenstein 1953/2008, Bruner 1983/1987, Clark 1996.

2 Herbert Clarks Ausdruck »common ground« wird in der sprachpsychologischen Literatur mit »gemeinsames Weltwissen« oder »gemeinsame Welterfahrung« wiedergegeben. Um gegenüber einem Wissen, das verschiedene Individuen unabhängig voneinander erworben haben, und einem Wissen, das sie sich durch gemeinschaftliche Tätigkeiten angeeignet haben, neutral zu bleiben, wurde »common ground« im folgenden mit »gemeinsamer Hintergrund« übersetzt. (A.d.Ü.)

ist eine absolut entscheidende Dimension aller menschlichen Kommunikation, einschließlich der sprachlichen mit all ihren (Personalpronomina) *er, sie* und *es*.

Von einem evolutionären Gesichtspunkt aus gesehen, ist der andere bemerkenswerte Aspekt dieses alltäglichen Beispiels einer menschlichen Zeigegeste seine prosoziale Motivation. Ich informiere Sie über die wahrscheinliche Gegenwart Ihres Exfreunds oder den Standort Ihres gestohlenen Fahrrads einfach deshalb, weil ich glaube, daß Sie diese Dinge wissen möchten. Im Tierreich ist eine derartige nützliche Mitteilung von Information äußerst selten, selbst bei unseren nächsten Verwandten unter den Primaten (in Kapitel 2 werden wir uns mit entsprechenden Beispielen befassen, etwa mit Warnschreien und auf Nahrung hinweisenden Rufen). Wenn beispielsweise ein wimmerndes Schimpansenjunges nach seiner Mutter sucht, haben mit an Sicherheit grenzender Wahrscheinlichkeit alle anderen Schimpansen in der unmittelbaren Umgebung davon Kenntnis. Wenn ein sich in der Nähe befindendes Weibchen nun weiß, wo die Mutter ist, wird sie es dem suchenden Jungen aber nicht sagen, obwohl sie durchaus in der Lage ist, ihren Arm in einer Art von Zeigegeste auszustrecken. Sie wird es dem Kind deshalb nicht sagen, weil es einfach nicht zu ihren Kommunikationsmotiven gehört, andere auf hilfreiche Weise über etwas zu informieren. Menschliche Kommunikationsmotive sind im Gegensatz dazu so grundlegend kooperativ angelegt, daß wir andere nicht nur über bestimmte Dinge auf hilfreiche Weise informieren, sondern wir einen Wunsch anderen einfach zur Kenntnis bringen, in der Erwartung, daß sie von sich aus Hilfe anbieten werden. So kann ich beispielsweise Wasser zum Trinken erbitten, indem ich einfach sage, daß ich es will (und Sie dadurch über meinen Wunsch informiere), wobei ich weiß, daß Ihre Neigung zu helfen (und unser gegenseitiges Wissen darüber) die Handlung des Informierens in den

meisten Fällen in eine vollständige Aufforderung verwandelt.

Menschliche Kommunikation ist somit ein grundlegend kooperatives Unternehmen, das am natürlichsten und reibungslosesten im Kontext eines wechselseitig vorausgesetzten, gemeinsamen begrifflichen Hintergrunds (1) und wechselseitig vorausgesetzter, kooperativer Kommunikationsmotive (2) funktioniert. Die grundlegend kooperative Eigenart menschlicher Kommunikation ist natürlich die zentrale Einsicht von Grice³ und wird in verschiedenen Graden und auf verschiedene Weise von anderen angenommen, die in dieser Tradition stehen, wie zum Beispiel Clark,⁴ Sperber und Wilson⁵ und Levinson.⁶ Wenn wir jedoch die letzten Ursprünge sowohl phylogenetisch als auch ontogenetisch verstehen wollen, müssen wir uns außerhalb der Kommunikation selbst umsehen und uns die menschliche Kooperation auf allgemeinere Weise anschauen. Es stellt sich heraus, daß die menschliche Kooperation in vielen Hinsichten einzigartig im Tierreich ist, sowohl was ihre Strukturen als auch ihre Motive betrifft.

Inbesondere wird die menschliche Kooperation durch etwas strukturiert, das einige zeitgenössische Handlungstheoretiker »geteilte Intentionalität« oder »Wir-Intentionalität« nennen.⁷ Im allgemeinen ist geteilte Intentionalität für die Beteiligung an spezifisch menschlichen Formen von Zusammenarbeit notwendig, bei denen ein Subjekt im Plural, ein »Wir« auftritt: gemeinsame Ziele, gemeinsame Absichten, wechselseitiges Wissen, geteilte Überzeugungen – und das alles im Kontext diverser Kooperationsmotive. Diese Gemein-

3 Grice 1957/1977, 1975/1979.

4 Clark 1992, 1996.

5 Sperber und Wilson 1986.

6 Levinson 1995, 2006.

7 Searle 1995/1997, Bratman 1992, Gilbert 1989.

samkeit ist besonders auffällig bei Interaktionen, die Institutionen betreffen und solche kulturell konstruierten Dinge wie Geld, Ehe und Regierung einschließen, die nur innerhalb einer institutionellen, kollektiv konstituierten Wirklichkeit existieren, an die wir alle glauben und in der wir gemeinsam handeln, so als ob es sie wirklich gäbe. Geteilte Intentionalität gibt es aber auch bei einfacheren und konkreteren Aktivitäten der Zusammenarbeit, etwa wenn wir uns das Ziel setzen, zusammen ein Werkzeug herzustellen oder zusammen spazierenzugehen, oder wenn wir einfach den Anblick eines Berges zusammen bewundern oder gemeinsam eine religiöse Handlung vollziehen. Der Vorschlag lautet also, die menschliche kooperative Kommunikation – gleichgültig, ob dabei »natürliche« Gesten oder »willkürliche« Konventionen benutzt werden – als einen Fall, wenn auch einen besonderen, einer nur beim Menschen vorkommenden Aktivität der Kooperation zu begreifen, die auf geteilter Intentionalität beruht.⁸ Die Kompetenzen und Motivationen geteilter Intentionalität machen somit das aus, was wir die »kooperative Infrastruktur der menschlichen Kommunikation« nennen können.

Wenn menschliche Kommunikation in einer Weise kooperativ strukturiert ist, wie das bei anderen Primaten nicht der Fall ist, stellt sich natürlich die Frage, wie sie entstanden sein könnte. Die Entstehung von Kooperation oder zumindest von Altruismus ist in der modernen Evolutionstheorie jedoch immer ein problematisches Thema. Wenn aber die Infrastruktur menschlicher kooperativer Kommunikation im Grunde dieselbe ist wie die aller anderen kollaborativen Aktivitäten, dann könnte es sein, daß sie sich als Teil einer umfassenderen menschlichen Anpassung in Richtung Kooperation und Kulturleben im allgemeinen entwickelte. Aus Gründen, die wir nicht kennen, hatten an einem bestimmten Punkt der

8 Tomasello, Carpenter, Call, Behne und Moll 2005.

menschlichen Entwicklung Individuen, die mit gemeinsamen Absichten, gemeinsamer Aufmerksamkeit und kooperativen Motiven ein gemeinsames Ziel verfolgen konnten, einen Anpassungsvorteil. Kooperative Kommunikation entstand dann als ein Mittel, diese Aktivitäten der Zusammenarbeit effizienter zu koordinieren, indem eine gemeinsame psychologische Infrastruktur geteilter Intentionalität zunächst vererbt und dann durch die Kommunikation weiter ausgebaut wurde. All das nahm mit ziemlicher Sicherheit seinen Anfang mit wechselseitigen Aktivitäten, bei denen ein Individuum, das seinem Partner half, zugleich sich selbst half. Dann fand jedoch eine Verallgemeinerung auf altruistischere Situationen statt, in denen Individuen einfach andere informierten oder mit ihnen zwanglos Dinge teilten, und zwar möglicherweise, um Reziprozität zu pflegen und zu demonstrieren, daß sie kooperativ sind. Erst später begannen die Menschen außerhalb kooperativer Kontexte auf diese neue kooperative Weise zu übergeordneten, nichtkooperativen Zwecken zu kommunizieren – bis sie sogar in der Lage waren, andere durch Lügen zu täuschen.

Die anfänglichen Schritte in diesem Prozeß vollzogen sich höchstwahrscheinlich im Modus der Geste. Das wird besonders deutlich, wenn wir die stimmliche und gestische Kommunikation unserer nächsten Verwandten unter den Primaten, den Menschenaffen, vergleichen. Die Vokalisierungen der Menschenaffen sind beinahe vollständig genetisch festgelegt, eng mit spezifischen Emotionen verknüpft und werden wahllos an alle in der unmittelbaren Umgebung gerichtet. Dagegen sind viele Gesten der Menschenaffen erlernt und werden in verschiedenen sozialen Situationen ganz flexibel zu verschiedenen sozialen Zwecken eingesetzt, wobei neue Gesten manchmal gelernt werden, um mit Menschen zu interagieren. Außerdem richten Menschenaffen diese Gesten an bestimmte Individuen und berücksichtigen dabei deren gegen-

wärtigen Aufmerksamkeitszustand. Lernen, Flexibilität und die Aufmerksamkeit für das Gegenüber sind offensichtlich grundlegende Merkmale der menschlichen Kommunikation. Bevor diese Merkmale existierten, konnten sich die Dinge nicht in die menschliche Richtung entwickeln. Wie viele Vertreter einer Theorie des gestischen Ursprungs zuvor bemerkt haben, ist es ebenfalls von Bedeutung, daß die menschliche Verwendung von Zeigegesten und Gebärdenspiel – als Nachfolger der Affengesten, sobald ein Zustand der Kooperation erreicht wurde – auf eine Weise »natürlich« ist, wie es arbiträre sprachliche Konventionen nicht sind. Insbesondere beruht das Zeigen auf der natürlichen Neigung von Menschen, der Blickrichtung von anderen zu externen Objekten zu folgen, und das Gebärdenspiel beruht auf der natürlichen Neigung, die Handlungen anderer als absichtlich zu interpretieren. Diese Natürlichkeit macht diese Gesten zu guten Kandidaten für einen Zwischenschritt in der Entwicklung der Kommunikation von Menschenaffen zu willkürlichen sprachlichen Konventionen.

Wie steht es nun mit der Sprache? Die vorliegende Hypothese macht geltend, daß willkürliche sprachliche Konventionen evolutionär nur innerhalb des Zusammenhangs von kollaborativen Aktivitäten, in denen die Teilnehmer Absichten und Aufmerksamkeit teilen und die durch natürliche Formen gestischer Kommunikation koordiniert werden, in Erscheinung treten konnten. Konventionelle Sprachen (die zuerst Zeichensprachen und dann stimmliche Sprachen waren) entstanden also auf dem Rücken dieser bereits verstandenen Gesten und ersetzen die Natürlichkeit des Zeigens und Gebärdenspiels durch eine gemeinsame Geschichte des sozialen Lernens (von der alle wechselseitig wissen, daß sie gemeinsam ist). Dieser Prozeß wurde natürlich durch die einzigartigen Fertigkeiten der Menschen zum kulturellen Lernen und zur Imitation ermöglicht, wodurch sie in die Lage versetzt

werden, von anderen und deren intentionalen Zuständen auf einzigartig leistungsfähige Weisen zu lernen.⁹ Im Rahmen derselben evolutionären Entwicklung begannen Menschen, verschiedene grammatikalische Konventionen zu schaffen und kulturell weiterzugeben, die in komplexe sprachliche Konstruktionen gegliedert wurden, welche komplexe *Typen* von Botschaften zum Gebrauch in wiederkehrenden Kommunikationssituationen festschrieben.

Wir benötigen also grundlegende evolutionäre Prozesse, die auf unterschiedliche Weisen operieren, um den Ursprung der zugrundeliegenden psychologischen Infrastruktur menschlicher kooperativer Kommunikation zu erklären. Um die Ursprünge der 6000 verschiedenen konventionellen menschlichen Sprachen zu erklären, brauchen wir zusätzlich noch kulturgeschichtliche Prozesse, durch die besondere sprachliche Formen in einzelnen Sprachgemeinschaften konventionalisiert werden. Abfolgen dieser Formen werden dann in grammatikalische Konstruktionen aufgenommen, und alle diese Konventionen und Konstruktionen werden schließlich durch kulturelles Lernen an neue Generationen weitergegeben. So können wir hier besonders deutlich die anhaltende Dialektik zwischen evolutionären und kulturgeschichtlichen Prozessen beobachten, die zuerst von Vygotskij¹⁰ und in einem moderneren evolutionären Rahmen von Richerson und Boyd¹¹ beschrieben wurde – und von der ich selbst eine Zeit lang besessen war.¹² Diese Perspektive auf die menschliche Kommunikation und Sprache stellt somit im Grunde den Vorschlag Chomskys vom Kopf auf die Füße, da die grundlegendsten Aspekte menschlicher Kommunikation als biolo-

9 Tomasello 1999/2002.

10 Vygotskij 1978/1992.

11 Richerson und Boyd 2005.

12 Tomasello, Kruger und Ratner 1993, Tomasello 2002, Tomasello et al. 2005.

gische Anpassungen zur Kooperation und sozialen Interaktion im allgemeinen aufgefaßt werden, während die im engeren Sinne sprachlichen Dimensionen der Sprache, einschließlich der grammatikalischen, kulturell konstruiert und von einzelnen Sprachgemeinschaften weitergegeben werden.

Der Weg zur modernen menschlichen Kommunikation war insgesamt betrachtet höchstwahrscheinlich ein sehr langer, ziemlich weitschweifiger und von vielen Wendungen durchsetzter. Um eine theoretische Darstellung zu bieten, die hauptsächlich auf empirischen Daten beruht, müssen wir daher viele verschiedene Aspekte des Lebens von Menschenaffen und Menschen betrachten – was diese Darstellung ebenfalls lang und weitschweifig ausfallen läßt. Aber trotz der vielen Verwicklungen, denen wir unterwegs begegnen, ist unser endgültiger Bestimmungsort eindeutig und läßt sich leicht angeben: die Identifikation der für die Spezies einzigartigen Merkmale menschlicher Kommunikation und ihre ontogenetischen und phylogenetischen Wurzeln. Im Hinblick auf dieses Ziel werde ich im folgenden drei spezifische Hypothesen bewerten:

1. Die menschliche kooperative Kommunikation trat in der Evolution zuerst (und tritt in der Ontogenese zuerst) in der Form natürlicher, spontaner Gesten des Zeigens und des Gebärdenspiels auf.

2. Die menschliche kooperative Kommunikation beruht wesentlich auf einer psychologischen Infrastruktur geteilter Intentionalität, die im Laufe der Evolution zur Unterstützung von Aktivitäten der Zusammenarbeit entstand und deren wichtigste Komponenten die folgenden sind: (a) sozio-kognitive Fertigkeiten zur gemeinschaftlichen Erzeugung gemeinsamer Absichten und gemeinsamer Aufmerksamkeit (und anderer Formen eines gemeinsamen begrifflichen Hintergrunds), und (b) prosoziale Motive (und sogar Normen) des Helfens und Teilens mit anderen.

3. Konventionelle Kommunikation, wie sie in menschlichen Sprachen verkörpert ist, wird nur möglich, wenn die an ihr Beteiligten schon über folgendes verfügen: (a) natürliche Gesten und ihre Infrastruktur geteilter Intentionalität sowie (b) Fertigkeiten des kulturellen Lernens und der Nachahmung, um gemeinsam verstandene kommunikative Konventionen und Konstruktionen zu schaffen und weitergeben zu können.