

## Von subhumaner zu humanspezifischer Interaktion: Tomasello und die Qualität des Zeigens

Rafael Mollenhauer, Universität Duisburg-Essen

**Summary.** Michael Tomasello, whose research is increasingly recognized in socio-theoretical discourses, describes deictic gestures as the crucial milestone on the way from animal interaction to specific human forms of interaction. On a mentalistic principle, he does not shy away from comparing the pointing gesture to language and from defining the deictic gesture as ontogenetic and phylogenetic earliest form of specific human cooperative communication. This article initially shows that the expression-theoretical construction of the research program creates contradictions and circularities. With reference to Karl Bühler's works, a semiotic analysis of Tomasello's examples of the special quality of pointing gestures is provided. This approach also casts a new light on Tomasello's explanation of the transition to verbal language and argues that Tomasello's theoretical construction hardly yields reliable statements according to social interaction as a process and to the emergence of cognitive capacities and communication means, but that his findings on the actual cognitive capacities of children and nonhuman primates could additionally be taken into account to unlock the secret of the origins of language.

**Zusammenfassung.** Die zunehmend auch in sozialtheoretischen Diskursen beachtete Forschung Michael Tomasellos erhebt deiktische Gesten zum entscheidenden Umschlagpunkt zwischen animalischen und spezifisch menschlichen Interaktionsformaten. Auf einer mentalistischen Grundlage scheut Tomasello nicht davor zurück, das Zeigen qualitativ gleichzusetzen mit der gesprochenen Sprache und es zur ontogenetisch und phylogenetisch frühesten Form humanspezifischer kooperativer Kommunikation zu erklären. In diesem Beitrag sollen zunächst aus der ausdrückstheoretischen Anlage des Forschungsprogramms resultierende Widersprüche und Zirkularitäten aufgedeckt werden, um anschließend unter Zuhilfenahme der Arbeiten Karl Bühlers eine funktionalistische Analyse der von Tomasello in verschiedenen Beispielen dargelegten deiktischen Qualität vorzunehmen, die auch ein neues Licht wirft auf Tomasellos Erklärung des Übergangs zur gesprochenen Sprache. Abschließend wird gezeigt, dass Tomasellos Ansatz kaum verlässliche Aussagen zur sozialen Praxis sowie zur Genese von kognitiven Vermögen und Kontaktmitteln generiert, dass seine Erkennt-

nisse zur kognitiven Leistungsfähigkeit von Kindern und nichtmenschlichen Primaten aber unterstützend hinzugezogen werden sollten, um den tatsächlichen Übergang zur menschlichen Sprache näher zu entschlüsseln.

## 1. Einleitung

Die Frage nach dem Ursprung der Sprache beschäftigt die Menschheit seit geraumer Zeit. In den vergangenen zwei Jahrzehnten war es vor allem der amerikanische Entwicklungspsychologe Michael Tomasello, der sich in diesem Kontext über die Zusammenführung seiner Untersuchungen zum kindlichen Spracherwerb, zur Ontogenese der Sozialkognition sowie zur Kognition nichtmenschlicher Primaten auch über die Grenzen der Psychologie hinaus einen Namen gemacht hat. Gerade mit seinem Werk *Die Ursprünge der menschlichen Kommunikation* (2009) offeriert Tomasello eine Art ‚Prototheorie der Handlungskoordination‘, die auch sozial- und handlungstheoretische Diskurse mittlerweile verstärkt beeinflusst. Im Zuge seines Erklärungsansatzes geht Tomasello davon aus, dass mittels deiktischer Gesten ebenso leistungsstark kommuniziert werden kann wie mit Hilfe der gesprochenen Sprache, dass also bereits die Verwendung einer (von geteilten Intentionen getragenen) Zeigegeste den Übergang von subhumanen zu humanspezifischen Interaktionsformen markiert. Doch bergen Tomasellos diesbezügliche Ausführungen ein großes Irritationspotential, das nachfolgend veranschaulicht werden soll (2), um anschließend das tatsächliche semiotische Potential der verschiedenen Kontaktmittel zu bestimmen (3) und auf dieser Basis auch Tomasellos Erläuterungen hinsichtlich des phylogenetischen Übergangs zur gesprochenen Sprache zu bewerten (4). Abschließend wird den Gründen einer beschränkten explikativen Reichweite des Ansatzes nachzugehen sein (5), nicht ohne den Versuch, einige Überlegungen Tomasellos für eine genauere Bestimmung des Übergangs zu humanspezifischen Kommunikationsformen fruchtbar zu machen (6).

## 2. Deiktische Gesten im Ansatz Tomasellos

Es sind deiktische Gesten, die Tomasello sowohl in seinem angenommenen evolutionären Szenario als auch mit Blick auf die kindliche Ontogenese zur grundlegendsten Form humanspezifischer kooperativer Kommunikation bestimmt. Getragen von geteilter Intentionalität sei die Zeigegeste der gesprochenen Sprache qualitativ ebenbürtig. Zur Unterfütterung dieser These zieht Tomasello in seinem kommunikationstheoretisch ambitionierten Werk *Die Ursprünge der menschlichen Kommunikation* immer wieder Beispiele heran, derer zwei hier als Ausgang der Argumentation dienen sollen. Beschrieben wird (1) eine Interaktion, die sich im hinteren Teil eines Flugzeugs abspielt: Während sich ein Mann in der Nähe der Flug-

zeugtoilette ein wenig ausstreckt, nähert sich eine Frau. Als sie die andere Person wahrnimmt, deutet sie mit fragendem Blick auf die Toilettentür, was nach Tomasello (2009: 75) wie folgt zu deuten sei: „Richte deine Aufmerksamkeit auf die Toilette; wartest du, daß sie frei wird?“ Weiterhin beschreibt Tomasello (2) eine Situation in einer Bar: Ein Mann, der noch etwas trinken möchte, wartet, bis der Barkeeper ihn ansieht, und zeigt daraufhin auf sein leeres Schnapsglas. Die Intention hinter der Zeigegeste lautet nach Tomasello (2009: 74): „Richte deine Aufmerksamkeit auf das leere Glas; fülle es bitte mit Schnaps.“

Die augenscheinliche Komplexität solcher Verwendungen deiktischer Gesten begründet Tomasello hier mit der Aufspaltung von referentieller und sozialer Intention. Nicht allein kompetente Sprecher seien in der Lage, Zeigegesten derart einzusetzen, denn schon Kleinkinder verwiesen mit Hilfe der Zeigegeste auf die unterschiedlichsten Entitäten, um unzählige soziale Intentionen auszudrücken, ohne dass sie bereits über sprachliche Fertigkeiten verfügten (Tomasello 2009: 76f.). Im Rahmen der zitierten Beispiele und hinsichtlich ihres Vergleichs mit den Zeigegesten von Kleinkindern verschweigt Tomasello jedoch, dass er neben dem Hinzutreten der sozialen zur referentiellen Intention auch den gemeinsamen begrifflichen Hintergrund zur notwendigen Voraussetzung der komplexesten Verwendungen deiktischer Gesten (und der Vielzahl möglicher sozialer Intentionen) erklärt. So bemüht er an anderer Stelle seines Werkes (Tomasello 2009: 13–16) eine Situation, in deren Rahmen Person A gegenüber Person B auf ein Fahrrad an der Mauer einer Bibliothek deutet. Je nach gemeinsamem Hintergrund könne diese Geste sehr verschiedene, komplexe Bedeutungen annehmen, zum Beispiel dann, wenn beide wissen, dass das Fahrrad dem Exfreund von A gehört, oder wenn beide sich gefragt haben, ob die Bibliothek zu jener späten Stunde noch geöffnet ist, sodass an der Wand lehrende Fahrräder einen Hinweis liefern könnten. Ein Beispiel, in dessen Kontext vorsprachliche Kinder ihre Verwendung der Zeigegeste auf einen so gearteten gemeinsamen begrifflichen Hintergrund stützen, kann Tomasello indes nicht bereitstellen, denn ein gemeinsamer begrifflicher Hintergrund müsste erst sprachlich (oder auf andere Weise symbolisch) konstituiert werden. Der Vergleich der zitierten Beispiele mit vorsprachlichen Zeigegesten konzentriert sich daher auf den Hinweis einer Aufspaltung von referentieller und sozialer Intention.

Verstärkt werden die Unklarheiten, sobald man sich dem von Tomasello aufgestellten Kooperationsmodell zuwendet. Hier bestimmt Tomasello (2009: 84) eine psychische Infrastruktur geteilter Intentionalität zur hinreichenden Voraussetzung spezifisch menschlicher Kommunikation. Neben sozialen Motivationen des Helfens und Teilens beinhalte diese Infrastruktur auch „kognitive[.] Fertigkeiten zur Erzeugung gemeinsamer Intentionen und Aufmerksamkeit (und anderer Formen eines gemeinsamen begrifflichen Hintergrunds) mit anderen“. Tomasello betrachtet somit (a) gemeinsame Aufmerksamkeit als Form eines gemeinsamen begrifflichen Hintergrunds und folglich beide als gleichrangig. Zudem sieht er sie (b) als bereits

in der vorsprachlichen Kognition verankerte Formate an. Doch werden hier entscheidende Erklärungsschritte übersprungen – und Qualitätsebenen vermischt: Berechtigterweise stellt Tomasello fest, dass Kinder mit etwa einem Jahr beginnen, an Aktivitäten gemeinsamer Aufmerksamkeit teilzunehmen und gemeinsame Ziele mit anderen zu verfolgen, und dass sie ab diesem Zeitpunkt auch Zeigegesten vermehrt im Rahmen dieser Aktivitäten einsetzen. Allerdings ist der gemeinsame begriffliche Hintergrund auf dieser Ebene der Handlungskoordination längst noch nicht gegeben. Zwar kann gemeinsame Aufmerksamkeit auch ausschließlich mit signalartigen Kontaktmitteln einhergehen, ein gemeinsamer begrifflicher Hintergrund, den Tomasello schon durch kognitive Fertigkeiten zur Erzeugung gemeinsamer Aufmerksamkeit und gemeinsamer Intentionen gesichert sieht, kann jedoch erst auf der Grundlage geteilter Sprachpraxis aufgebaut werden. Ohne sprachlich kommuniziert zu haben, lässt sich diese Ressource folglich gar nicht nutzen. Die Beteiligten in Tomasellos Fahrrad-Beispiel müssen also bereits auf symbolischer Ebene kommuniziert haben über den Exfreund von A oder die Möglichkeit einer zu dieser späten Stunde bereits geschlossenen Bibliothek. Im beschriebenen Flugzeug-Beispiel stützen sich die Beteiligten dagegen weniger auf ein spezifisch von ihnen konstituiertes gemeinsames Wissen als auf geteilte gesellschaftliche Normen und, wie auch im Bar-Beispiel, auf geteiltes Wissen hinsichtlich gängiger Handlungszusammenhänge. Es mag möglich sein, einen akkulturierten Affen darauf zu trainieren, in einer Bar (empraktisch) auf ein leeres Glas zu deuten, um etwas zu trinken zu bekommen, dies ändert jedoch nichts an einem auch in Tomasellos vergleichsweise einfach gehaltenen Bar-Beispiel – zumindest die Institution ‚Bar‘ betreffend – symbolisch konstituierten Handlungszusammenhang. Ein wie auch immer geartetes symbolisches Hintergrundwissen bestimmt aber nicht die Qualität des Zeigens selbst – gerade deshalb kann der Affe ja zeigen, ohne am gemeinsamen begrifflichen Hintergrundwissen zu partizipieren.

Tomasello bestimmt das Zeigen demgegenüber zur Urform humaner Kommunikation, indem er Fälle anführt, die nur auf der Grundlage bereits vollzogener symbolischer Kommunikationsakte denkbar sind, und überträgt diese Fälle dann auf kindliche Verwendungen der Zeigegeste, indem er den gemeinsamen begrifflichen Hintergrund qualitativ gleichsetzt mit gemeinsamer Aufmerksamkeit. Eng verknüpft sind die beschriebenen Unklarheiten und Widersprüche damit, dass Tomasello kaum in den Blick nimmt, was die einzelnen Kontaktmittel selbst in der gegenseitigen Verhaltensabstimmung wirklich vermögen. Die Zeigegeste leistet in allen angeführten Beispielen nämlich allein die Markierung eines Gegenstandes in einer gegenwärtigen Situation gemeinsamer Wahrnehmung (im Hier und Jetzt) – ganz so, wie es zumindest auch von akkulturierten Affen zu erwarten ist. Sie mag in der von Tomasello beschriebenen Art und Weise Verwendung finden, um auf komplexe Sachverhalte zu verweisen, aus eigener Kraft entfaltet die allenfalls empraktisch (siehe unten) wirkende Zeigegeste jedoch keine neue Sinndimension. Es ist demnach analytisch zu

unterscheiden zwischen (a) der Frage nach der Leistungsfähigkeit des Zeigens selbst und (b) der Rolle eines gemeinsamen begrifflichen Hintergrunds.

Bei Tomasello, der diese Ebenen miteinander vermischt, zieht sich der Versuch des Nachweises einer qualitativen Gleichrangigkeit von Zeigegesten und sprachlichen Kommunikationsakten indes durch das gesamte Werk. Immerhin setzt Tomasello das Zeigen nicht in allen Belangen gleich mit der gesprochenen Sprache. So seien „Zeigegesten [...] notwendig wirkungslos in Situationen, in denen die Beteiligten nur wenig oder überhaupt keinen gemeinsamen Hintergrund haben, besonders dann, wenn lange Ketten von Schlussfolgerungen erforderlich sind.“ (Tomasello 2009: 217) Während Tomasello (2009: 12f.) zuvor noch auf die Leistungsfähigkeit der Zeigegeste auch in einer fremden Kultur mit einer fremden Sprache verwiesen hatte, gesteht er nun ein, dass die Zeigegeste aufgrund ihrer Abhängigkeit vom gemeinsamen Hintergrund der Beteiligten auch keine wirksame Kommunikationsform unter Fremden und zum Unterrichten von Kindern sei. Doch sieht er in dieser Abhängigkeit nicht allein eine Schwäche, sondern auch eine Stärke. Im Grunde werde der gemeinsame Hintergrund im Rahmen sprachlicher Kommunikation lediglich durch eine gemeinsame Geschichte sozialen Lernens ersetzt. Auch diese These unterfüttert Tomasello (2009: 217) mit einem Beispiel, das allerdings nur zusätzliche Verwirrung stiftet. Imaginiert wird eine Situation, in der A und B auf gemeinsame Jagderfahrung an einem Wasserloch zurückblicken und an besagtem Ort mehrfach eine Gazelle gesichtet haben. Kommt Akteur A nun aufgeregt aus der Richtung des Wasserlochs angelaufen und zeigt in eben diese Richtung, könne B annehmen, dass A dort abermals eine Gazelle gesichtet hat. Ohne geteilte Erfahrung in vorangegangenen Situationen könne A hingegen „natürlich nicht auf den abwesenden Bezugsgegenstand zeigen.“ (Tomasello 2009: 217) Diese (angebliche) Möglichkeit der Bezugnahme auf Abwesendes deutet Tomasello (2009: 124–129) in seinem kognitiv reichhaltigen Modell menschlicher Zeigegesten als Beleg einer Kommunikation auf der Ebene kognitiv repräsentierter Entitäten. Allerdings könnte die Zeigegeste sehr wohl in genau der beschriebenen Weise Verwendung finden, wenn A und B keinerlei gemeinsame Jagderfahrung am Wasserloch aufwiesen. Dann würde B die Zeigegeste zwar nicht mit einer spezifischen Erfahrung am Wasserloch in Verbindung bringen, könnte die Situation aufgrund bestimmter anderer Erfahrungen aber dennoch auf seine Art deuten – auch weil die Zeigegeste, wie nachstehend näher erläutert wird, die außerhalb gemeinsamer Wahrnehmung befindliche Entität (anders als von Tomasello angenommen) gar nicht trifft.

### **3. Eine semiotische Aufschlüsselung**

Eine genauere Aufschlüsselung der tatsächlichen Qualität der unterschiedlichen Kontaktmittel versprechen die Arbeiten Karl Bühlers (1999, 2000),

der darauf hinweist, dass mit der Zeigegeste nur Entitäten in einer gegenwärtigen Situation gemeinsamer Wahrnehmung getroffen werden können (Bühler 1999: 168). Sie eignet sich weder, um auf etwas derzeit sinnlich nicht Wahrnehmbares zu deuten, noch kommt ihr in Abwesenheit eines Gegenübers Bedeutung zu. Wenn Tomasello der Zeigegeste also informierenden Charakter auch hinsichtlich räumlich und zeitlich nicht gegenwärtiger Dinge und Sachverhalte zugesteht, begeht er insofern einen Fehlschluss, als die Zeigegeste selbst diese Dinge und Sachverhalte nicht erreicht. In Tomasellos Wasserloch-Beispiel zeigt der Akteur nicht auf die Gazelle, die man aus gemeinsamer Erfahrung kennt, sondern allein in die Richtung des Wasserlochs, an dem man der Gazelle bereits gemeinsam begegnet ist (der Verweis auf die Gazelle ergibt sich aus dem Handlungszusammenhang). Dass ähnlich geartete Formen der Zeichenverwendung schon bei vergleichsweise einfachen Lebewesen anzutreffen sind, veranschaulicht Bühlers Beispiel der Honigbienen, die ihrem Gegenüber durch einen Tanz Richtung und Entfernung von Blüten anzeigen und ihren Artgenossen zudem eine Stoffprobe offerieren (Bühler 2000: 72–74). Einen derartigen Fall deckt Bühler schon im Rahmen seines ersten Axioms aus der *Krise der Psychologie* ab, wenn er schreibt, spezifisch semantische Einrichtungen müssten zum Einsatz kommen, sobald die Richtpunkte der Steuerung nicht in gemeinsamer Wahrnehmung gegeben sind (Bühler 2000: 71). Solche Einrichtungen betreffen, wie das Beispiel der Honigbienen zeigt, nicht notwendigerweise entindexalisierte Kontaktmittel, die entstofflicht und ablösbar von den Bezugsentitäten sind. Nur weil ein Steuerungsrichtpunkt den Beteiligten nicht in der gemeinsamen Wahrnehmungssituation direkt zugänglich ist, müssen also nicht zwingend sprachliche beziehungsweise symbolische Kontaktmittel zum Einsatz kommen.

Die spezifische Semantizität der menschlichen Sprache (das vollends entfaltete Symbol) leitet Gerold Ungeheuer (2004) in seinem Aufsatz zur kybernetischen Grundlage der Sprachtheorie Bühlers in mehreren Schritten aus der Grundsituation menschlichen Verhaltens ab (siehe auch Mollenhauer 2010: 28–33). Dabei wird auch die Annahme als Trugschluss entlarvt, die symbolische Ebene, wie sie mit dem Perspektivenwechsel im dritten Axiom der *Krise der Psychologie* beschrieben wird, sei bereits mit der Verwendung eines einzelnen Nennwortes erreicht. Auch das alleinstehende Wort kann für sich genommen nur symphysisch oder empraktisch (bzw. sympraktisch) zum Einsatz kommen (siehe auch Bühler 1999: 155–161), seine Verwendung bedeutet also noch kein Überschreiten signalhafter Kontakte. Sowohl der symphysische als auch der empraktische Einsatz der Zeichen können aber die Grundsituation gemeinsamer Wahrnehmung, welche die Einstellung der Individuen aufeinander als notwendige Voraussetzung des Steuerungskontakts ebenso beinhaltet wie das zur Setzung der Richtpunkte als Steuerungsorientierung erforderliche gegenseitige Verstehen der Tätigkeiten des anderen, schon überschreiten. Im Vergleich mit einer (im animalischen und menschlichen Bereich vorzufindenden) gegenseitigen Steuerung in gemeinsamer Wahrnehmungssituation, die zwar als



elementare Basis auch des sprachlichen Vollzugs anzusehen ist (Ungeheuer 2004: 131), dennoch aber nur ein Minimum semantischer Einrichtungen verlangt, muss im Hinblick auf symphysische und empraktische Zeichenverwendungen von einer Verbesserung der Kommunikationsmittel die Rede sein, die erforderlich wird, wenn den Beteiligten nur Teile oder Abschnitte des Aktionsschauplatzes sinnlich zugänglich sind. Trotzdem erschöpfen beide Formen der Zeichenverwendung sich aber in den Funktionen des Benennens und Verweisens.

Ein symphysischer Gebrauch der Kommunikationszeichen liegt vor, wenn erst das physische Umfeld die Grundlage zur Interpretation der Zeichen liefert, die Zeichen abgelöst von ihm also kommunikativ und sozial sinnlos wären. Zu differenzieren sind hier die Kategorien ‚Stoffgebundenheit‘ und ‚Gegenstandsgebundenheit‘. Stoffgebunden sind beispielsweise die oben genannten Blütenproben der Honigbienen, Gegenstandsgebundenheit liegt im Falle eines durch einen Hund markierten Baumes ebenso vor wie im Falle eines Markennamens auf einer Nahrungsmittelverpackung. Entscheidend ist, dass das Zeichen seine Identifikationsfunktion nicht abgelöst vom markierten Gegenstand erfüllen kann. Geprägt von derart symphysischen Zeichenverwendungen ist auch die sogenannte Affensprachforschung, die Tomasellos kognitionstheoretisch fundierten Ansatz immer wieder vor Erklärungsprobleme stellt, während Bühlers Theorieangebot und dessen Interpretation durch Ungeheuer die Beschränkungen der Tiere bezüglich symbolischer Akte schnell aufdecken (Mollenhauer 2010).

Im Hinblick auf den empraktischen Einsatz der Kommunikationszeichen bemüht Bühler (1999: 155) ein Beispiel, das dem von Tomasello angeführten Bar-Beispiel recht ähnlich ist. Während der Gast in Tomasellos Beispiel auf das beiden Beteiligten sinnlich wahrnehmbare Glas zeigt, beschreibt Bühler einen wortkargen Gast, der im Kaffeehaus „einen schwarzen“ bestellt. Das intendierte Getränk kann hier gegenwärtig sinnlich wahrnehmbar sein oder nicht, es könnte sich also beispielsweise auch unter der Theke oder im Hinterzimmer und damit nicht im Sichtfeld des Gastes befinden. Vorstellbar ist auch, dass der Gast auf die Tür zum Hinterzimmer des Kaffeehauses zeigt, folglich also wie in Tomasellos Gazellen-Beispiel nicht den intendierten Gegenstand selbst trifft, sondern stattdessen die Richtung, in der sich der Gegenstand befinden soll. In all jenen Fällen kann sowohl die Zeigegeste als auch das einzelne Wort, ob die zu bezeichnende Entität direkt zugegen ist oder sich als Steuerungsrichtpunkt außerhalb gemeinsamer Wahrnehmung befindet, nur über den Handlungszusammenhang gedeutet werden. Dies ist der empraktische Einsatz der Kommunikationszeichen, den man im Tierreich kaum antrifft. Zumindest bei akkulturierten Affen sind derartige Formen der Zeichenverwendung aber möglich. Der Großteil der von Tomasello mit Blick auf die besondere Qualität deiktischer Gesten angeführten Zeichenverwendungen ist derart empraktischer Art. So wie das einzelne Wort stets eine schlichte Benennung darstellt, bleibt die Zeigegeste aber immer ein bloßer Verweis, in welchem Maße auch immer der Handlungszusammenhang begrifflich bzw. symbolisch konstituiert ist.

Tomasello bleibt in all seinen Beispielen auf einen empraktischen Einsatz der Zeigegeste beschränkt. Zudem thematisiert er kaum, dass die zwischenmenschliche Kommunikation in erheblichem Maße durch signalartige Kontaktmittel – darunter symphysische und empraktische Zeichenverwendungen – strukturiert wird. Der Grund hierfür liegt vermutlich in der Nichtberücksichtigung nebeneinander existierender Ebenen von Intentionalität. Grundlage einer Kontaktmittelqualität ist in Tomasellos Ansatz stets ein bestimmtes Niveau intentionaler Tätigkeiten; so ist die nach Tomasellos Auffassung bereits humanspezifische Zeigegeste durch kognitive Fertigkeiten und Motivationen geteilter Intentionalität bestimmt. Sind humanspezifische kognitive Vermögen (Fertigkeiten geteilter Intentionalität) und die mutmaßlich davon abhängigen Kommunikationsmittel (deiktische Gesten) erst einmal abrufbar, scheinen sie aus Tomasellos Sicht die gesamte artspezifische soziale Praxis zu bestimmen. Zumindest wird kaum thematisiert, inwieweit Menschen auch auf Ebenen interagieren, die in qualitativ vergleichbarer Weise auch im Tierreich vorzufinden sind. Die wirkliche semantische Qualität und die spezifischen Eigenschaften der semiotischen Mittel erfasst Tomasello auch deshalb nicht, weil er sie nicht im Kontext der gegenseitigen Verhaltensabstimmung betrachtet. Im Gegensatz zu Bühler beginnt er nicht bei der Gemeinschaft und den dort signalartig ablaufenden Kontakten, sondern bei einem Individuum und seinen jeweils artspezifischen (kognitiv-motivationalen) Vermögen. Unberücksichtigt bleibt daher auch, was die menschliche Sprache über die Leistungsfähigkeit der von Tomasello zum Ursprung humanspezifischer Kontakte erhobenen Zeigegeste hinaus zu leisten vermag. Stehen nicht (mindestens) zwei in gegenseitige Steuerungsprozesse verwickelte Interaktionspartner am Anfang der Überlegungen, sind auch die von Bühler behandelten Richtpunkte der Steuerung nicht vollends definiert. Damit entfällt die Möglichkeit, Kontaktmittel hinsichtlich ihrer Situationsgebundenheit oder Situationsentbundenheit zu bewerten. Gerade die von Bühler herausgestellte Darstellungsebene der Sprache – die spezifisch sprachliche Semantizität – ist aber über ihre Situationsentbundenheit definiert. Allein im Prozess der Handlungskoordination entfaltet sich das vom Hier und Jetzt (und damit auch vom eigenen Körper und der eigenen Befindlichkeit) ablösbare und entstofflichte Symbol als neue und eigenständige Sinndimension. Auch Symbole bleiben natürlich verknüpft mit einem den beteiligten Individuen zugänglichen Weltausschnitt, in dessen Kontext die Interaktanten sich der Zeichen bedienen, das Sinnverstehen ist aber nicht mehr abhängig vom jeweiligen Handlungsgefüge. Der stoffliche Träger besteht sowohl beim Schallphänomen gesprochener Sprache als auch bei einem auf Papier geschriebenen Text, es ist aber nicht dieses Umfeld, sondern vielmehr das synsemantische Umfeld, das die Bedeutung der Zeichen bestimmt und das Symbol zur vollen Entfaltung gelangen lässt.

Bei einem synsemantischen Einsatz der Kommunikationsmittel, den Bühler als Steigerung des symphysischen und empraktischen Kontakts und als Spiegel spezifisch sprachlicher Semantizität einführt, wird die Bedeu-



tung eines Wortes erst über die das Wort umgebende Rede bestimmt. Das kraft Konvention eine bestimmte Bedeutung vermittelnde Symbol erfüllt seine Funktion demnach nur innerhalb des Symbolfeldes. Zur Veranschaulichung mag hier Bühlers *Phantasma* als ein kommunikativer Entwurf eines Handlungsraumes im Bewusstsein der Kommunikationspartner dienen. Mit Hilfe sprachlicher Mittel werden die Vorstellungen des Gegenübers derart gesteuert, dass ihm die Szene der Aktion als Phantasma gegenwärtig wird (Ungeheuer 2004: 143f.). Als rein sprachlich aufgebautes Konstrukt enthält das Phantasma die Richtpunkte der Steuerung, sodass auch Akte des Verweisens innerhalb des Phantasmas stattfinden (die *Deixis am Phantasma*). In Tomasellos Beispielen zur (angeblich) humanspezifischen Qualität deiktischer Gesten ist ein derartig von den Bezügen des gegenwärtigen Handlungszusammenhangs losgelöstes Phantasma nie gegeben. Entweder befinden sich alle Richtpunkte der Steuerung in gegenwärtiger gemeinsamer Wahrnehmung (Bar-Beispiel) oder die Beteiligten selbst agieren in gegenwärtiger gemeinsamer Wahrnehmung als Steuerungsrichtpunkte und verweisen in die Richtung eines außerhalb gemeinsamer Wahrnehmung befindlichen Richtpunktes, der über den jeweiligen Handlungszusammenhang spezifiziert werden kann (Gazellen-Beispiel).

Auch den Weg zu einem gemeinsamen begrifflichen Hintergrund ebnet letztlich erst das synsemantische Potential der Sprache. Die Bedeutung der Sprachzeichen ist dabei keinesfalls durch die mit ihnen verbundenen Gedanken gesichert, sondern vielmehr in Abhängigkeit von den in sozialen Kontexten geltenden Konventionen zu sehen. Ihre besondere Leistungsfähigkeit gewinnt die Sprache also gerade dadurch, dass sie hinsichtlich ihrer Formate nicht mit denen der Kognition übereinstimmt (Loenhoff und Mollenhauer 2016: 123f.). Tomasello überspringt hier Erklärungsschritte, mit denen sich die Genese eines geteilten begrifflichen Hintergrunds erst beschreiben ließe, und bettet die Erfüllungsbedingungen der Sinnproduktion stattdessen ein in die sozio-kognitive Infrastruktur der Individuen, wodurch die Selbstausslegungen der Individuen gegenüber den gelingenden Handlungsanschlüssen primären Charakter annehmen, obgleich die kognitiven Vermögen dem Kommunikationsgeschehen tatsächlich erst erwachsen. Die Beziehung von gemeinsamem Hintergrund, kooperativer Kommunikation und individuellen Vermögen ist folglich erklärungsbedürftiger als Tomasellos evolutionsbiologisches Modell voraussetzungsloser Ausgangspunkte vermuten lässt (Loenhoff und Mollenhauer 2016: 122f.).

Das semiotische Potential gesprochener Sprache darf dennoch nicht darüber hinwegtäuschen, dass die Sprache nicht nur auf synsemantischer Ebene agiert, sondern dass sie stattdessen eine ökonomische Lösung darstellt, die auf unterschiedlichen Qualitätsebenen wirksam wird. Allein das Vorkommen sprachlicher Mittel bedeutet also nicht zwingend eine dem Menschen vorbehaltene Interaktionsqualität. Häufig funktionieren sprachliche Signale auch als einfache Kontaktmittel, die in ihrer sprachlichen Bedeutung abgeschwächt oder gänzlich entleert sind (Ungeheuer 2004: 138). Im Widerspruch dazu sieht Tomasellos intentionalistische Leiter kei-

nen ‚Rückschritt‘ im Sinne der Verwendung auch im subhumanen Bereich anzutreffender Kontaktmittelqualitäten durch humane Akteure vor. Ist die Ebene humanspezifischer Kommunikation in zunächst gestischer Modalität erst einmal erreicht, erscheinen demzufolge alle Steuerungsaktivitäten eines kognitiv entsprechend ausgerüsteten Individuums als spezifisch menschlich.

#### 4. Tomasello und der Übergang zur gesprochenen Sprache

Der von Tomasello angenommene Übergang zur gesprochenen Sprache beschreibt also einen Wechsel zu einer weniger auf den gemeinsamen Hintergrund als vielmehr auf eine gemeinsame Geschichte sozialen Lernens gestützte Kommunikationsform, nicht aber einen echten Qualitätssprung. Die mangelnde zeichentheoretische Reife dieser These verdeutlicht Tomasello (2009: 242–252) ungewollt noch einmal in seinen den ‚Wechsel zur stimmlichen Modalität‘ betreffenden Ausführungen. Seiner Auffassung nach „konnten die ersten Kommunikationskonventionen überhaupt nicht im Zusammenhang mit der stimmlichen Modalität entstanden sein“ (Tomasello 2009: 242), da (a) nichtmenschliche Primaten nicht absichtsvoll und ausschließlich in enger Bindung an ihre Emotionen vokalisiert und (b) Menschen zwar irgendwann willentliche Kontrolle über ihre Vokalisierungen erlangt haben müssen, diese aber im Gegensatz zu handlungsbasierten Gesten kein gutes Medium für die referentielle Intention darstellten. Es sei weder natürlich, (b1) die Aufmerksamkeit anderer mittels Vokalisierungen auf externe Gegenstände zu lenken, noch eigneten sich nichtkonventionalisierte Vokalisierungen, um (b2) die Vorstellungskraft anderer auf abwesende Entitäten zu lenken. Doch werden die genannten Punkte, soweit sie von der Zeigegeste (welche die gegenwärtige Situation gemeinsamer Wahrnehmung nicht transzendiert) erfüllt werden, durchaus auch von Steuerungsprozessen im Tierreich abgedeckt (zum Beispiel im Falle der Honigbienen). Der Nachweis einer zunächst gestisch ablaufenden Konventionalisierung ist selbst auf der Grundlage der Annahme einer besonderen Wirksamkeit von Gesten in derartigen Situationen also längst nicht erbracht. Auf lautlicher Ebene ist außerdem das einzelne Wort zweifelsfrei für einen derartigen Einsatz geeignet, wie auch Bühlers Beispiel des wortkargen Kaffeehausgastes veranschaulicht. Doch sucht Tomasello wohl nach Beispielen eines entsprechenden Einsatzes nichtkonventionalisierter Laute.

Nimmt man nun Bühlers Ausführungen aus der *Krise der Psychologie* zur Hand, erfüllt Tomasellos eigenes Beispiel der Warnrufe Grüner Meerkatzen sehr wohl die Bedingungen der ersten beiden Axiome: Sie werden innerhalb einer Gemeinschaft hervorgebracht, sie bringen die innere Befindlichkeit zum Ausdruck und sie dienen der Steuerung anderer Individuen. Da sie allerdings nicht intentional hervorgebracht würden und nicht willentlich an andere Individuen gerichtet seien, gelten sie Tomasello nicht als geeignetes Beispiel. Doch spielt das Maß der Intentionalität hier nicht die

entscheidende Rolle, wenn funktionalistisch die Wirksamkeit der semiotischen Mittel betrachtet wird. Ohne Zweifel kann der Grad der Intentionalität im Kontext signalartiger Steuerungen – von der Koordination mit Hilfe von Duftstoffen innerhalb eines Ameisenstaates über die gemeinsame Jagd der Wölfe bis hin zu signalartigen Kontakten unter Menschen – erheblich variieren (von schlichter Zweckgerichtetheit bis zur bewusst reflektierten Absicht mit Erwartungserwartungen), die Funktionalität der Zeichen im Steuerungsprozess ist allein hierdurch aber nicht bestimmt. Weder das bloße Vorhandensein intentionaler Phänomene noch eine graduelle Abstufung von Intentionalität soll damit gelehnet sein. Vielmehr geht es darum, die Hypothesen nicht durch spezifische (bei Tomasello mentalistische) Vorentscheidungen derart zu belasten, dass bestimmte Phänomene (hier der eigentliche Steuerungsprozess) aus dem Blick geraten. Bühler begründet daher seine Axiomatik in methodologischer Offenheit mit dem Behaviorismus (1. Axiom), dem er dann die Erlebnispsychologie hinzugesellt (2. Axiom). Dies gestattet sodann die Untersuchung der Genese mentaler Phänomene im Steuerungsprozess. Auf der Grundlage einer solchen Betrachtungsweise kann den Warnrufen der Affen schließlich sehr wohl das von Tomasello geforderte referentielle Potential zugesprochen werden. Immerhin bemerkt Tomasello (2009: 26f.) selbst, dass die Warnrufe Grüner Meerkatzen sich auf ganz bestimmte Dinge beziehen, der Schlangenwarnruf also vom Adlerwarnruf abgegrenzt ist, und dass außerdem die Adressaten – allein mit diesem Terminus widerspricht Tomasello schon seiner Hypothese individualistischer Ausdrücke von Emotion – entsprechende Informationen aus den Warnrufen extrahieren. Dass alle Affen innerhalb einer Art dasselbe grundlegende Repertoire von Rufen aufweisen, zeigt zudem eine durchaus spezifische (artspezifische) Appellfunktion an, die auf der Basis mentalistisch-ausdruckstheoretischer Vorentscheidungen kaum angemessen gewürdigt werden kann.

Seine Zweifel am referentiellen Potential nichtkonventionalisierter Vokalisierung verbindet Tomasello mit der Folgerung, dass es „im allgemeinen sehr schwierig ist, sich die ausschließlich in der stimmlichen Modalität stattfindende Entwicklung sinnvoller, menschenähnlicher kooperativer Kommunikation – geschweige denn von Kommunikationskonventionen – auch nur vorzustellen.“ (Tomasello 2009: 245; Hervorhebung R.M.) Im Bereich des Handelns, das Tomasello zur ursprünglichsten Form der Sinnhaftigkeit einer intentional bestimmten menschlichen Kommunikation erklärt, könne man sich jene Entwicklung aber leicht vorstellen. Ganz wesentlich beruhe die handlungsbasierte Infrastruktur menschlicher Kommunikation auf den humanspezifischen Veranlagungen, der Blickrichtung anderer automatisch zu folgen, andere durch Zeigen zum Folgen der eigenen Blickrichtung zu bewegen und die Handlungen anderer als intentional aufzufassen. Wenn Tomasello von der Ausgeschlossenheit einer sich ausschließlich im stimmlichen Modus vollziehenden Entwicklung humanspezifischer Kommunikation spricht, scheint er jedoch nebeneinander existierende Modalitäten zu veranschlagen, die in der Art einzelner Kanäle unabhängig von-

einander geöffnet oder geschlossen werden können. Natürlich wird die Genese humanspezifischer Kontakte nicht allein stimmlich vonstattengegangen sein, da Vokalisierungen selbstverständlich an weitere Realisierungsbedingungen gebunden sind. In vergleichbarer Weise gilt dies aber auch für die Zeigegeste. Bühler berücksichtigt eben diesen Tatbestand, wenn er spezifisch menschliche Kommunikationsakte in einfachsten Steuerungsprozessen innerhalb der Gemeinschaft fundiert.

Indem Tomasello alle echten Qualitätsmerkmale der gesprochenen Sprache auch dem Zeigen zugesteht – selbst die Artikulation von Propositionen scheint er der Zeigegeste aufgrund der angenommenen Bindung an den gemeinsamen (begrifflichen) Hintergrund zuzusprechen –, schafft er die Notwendigkeit, den Übergang zur gesprochenen Sprache anderweitig zu begründen. Unterschiede zwischen beiden Steuerungsmitteln spricht er durchaus an – so zum Beispiel die Abhängigkeit von einem gemeinsamen Hintergrund (Zeigen) gegenüber der Abhängigkeit von einer gemeinsamen Geschichte sozialen Lernens (Sprache) oder das größere referentielle Potential der Zeigegeste in gemeinsamer Wahrnehmung. Und in der Tat sind hinsichtlich der grundsätzlichen Wirksamkeit von Zeigegesten und Vokalisationen Differenzierungen notwendig: So funktioniert die Zeigegeste zum Beispiel stärker diakritisch, während die vokale Geste auch von ihrem Verwender selbst gehört wird, was ihr im Ansatz Meads (1973: 100–107) besondere Relevanz für den Übergang zum signifikanten Symbol und somit der menschlichen Sprache zuträgt und darüber hinaus (neben den Aktivitäten des Gegenübers) entscheidend das Verfertigen der Gedanken beim Reden (Kleist 1984) beeinflusst. Tomasello betont demgegenüber, dass Vokalisationen im Gegensatz zu Zeigegesten und Gebärden auch von anderen Individuen in der Umgebung gehört werden, demnach also an ein öffentliches Publikum gerichtet sind (2009: 246–248). Auch hier bestimmt zwar die kommunikative Absicht den Adressaten einer Zeichenverwendung, stimmliche Akte erlaubten anderen Individuen aber das Lauschen und seien durch ihre Öffentlichkeit relevant für die Reputation. In eben jener Öffentlichkeit vokaler Akte sei also die Grundlage für die heutige Dominanz der gesprochenen Sprache zu sehen.

Die frühesten stimmlichen Konventionen betrachtet Tomasello schlicht als emotionale Begleiterscheinungen handlungsbasierter Gesten; erst mit dem Auftreten einer stärkeren willentlichen Kontrolle über Vokalisierungen und nach dem Entstehen der kommunikativen Absicht hätten sie für sich einen funktionalen Charakter erlangt. Gleich in mehrfacher Hinsicht kann diese Sichtweise allerdings zum Spiegel eines mentalistischen Reduktionismus' erklärt werden: Zum einen entfällt schon die (augenscheinliche) Notwendigkeit, einen Übergang zu einer qualitativ vergleichbaren Form gegenseitiger Steuerung erklären und begründen zu müssen, wenn man das von Tomasello aufgrund entsprechender Vorentscheidungen weitgehend unentdeckte semiotische Potential der Sprache in den Blick nimmt. Darüber hinaus irritiert die Idee, eine stimmliche Konvention stelle sich schon durch einen (begleitenden) Ausdruck von Emo-

tion ein, denn erst die überindividuelle Ebene kennzeichnet konventionalisierte Kommunikation als solche. Auch das Auftreten einer stärkeren willentlichen Kontrolle und die Grice'sche kommunikative Absicht scheinen schließlich nur sehr allgemeine individuelle Voraussetzungen zu sein, die mit der eigentlichen Konventionalität symbolischer Mittel recht wenig gemein haben. Mit der Auffassung einer weitgehend individuell getragenen Sprache stützt Tomasello aber sein eigenes Modell einer bereits im gestischen Modus in ähnlicher Weise intentional getragenen humanspezifischen Kommunikation.

Auch die syntaktische Ebene von Äußerungen bettet Tomasello letztlich immer – wie schon motivational bestimmte Grammatiken des Aufforderns, Informierens und Teilens zur Schau stellen – in das Individuum ein (siehe insbesondere Tomasello 2003). Der artikulierte Satz stellt damit kaum mehr dar als einen Ausdruck kognitiver Vermögen bzw. Kompetenzen. Auch die mit der psychischen Infrastruktur geteilter Intentionalität verschränkten prosozialen Motive des Helfens und Teilens erheben die Zeigegeste natürlich nicht zu einem synsemantischen Kontaktmittel. Wie schon die kognitiven Vermögen entbinden sie die deiktischen Mittel weder von der gegenwärtigen Situation gemeinsamer Wahrnehmung noch determinieren sie die Qualität eines Kontaktmittels allein aufgrund ihrer Beteiligung oder ihres grundsätzlichen Vorhandenseins. Wenn Tomasello (2009: 19) Kooperativität in ihrer Abhängigkeit von den genannten Motiven als ‚humanspezifisch‘ betitelt, um auf dieser Grundlage in Form des Lügens und der Täuschung eine kooperative Form der Kommunikation zu nichtkooperativen Zwecken etablieren zu wollen, ist dies nur als weiterer Hinweis auf die Notwendigkeit eines sozialen Fundaments zu deuten. Mit Recht stellt er später fest, dass auch Lügen und Täuschung kooperative Kommunikation erfordern (Tomasello 2009: 205). Doch sollten hier nicht individualistische Motive, sondern vielmehr der intrinsisch kooperative Charakter eines jedweden Steuerungsprozesses im Vordergrund stehen. Auch altruistischen Motiven kann im Übrigen nur dann Bedeutung zukommen, wenn sie von einem Interaktionspartner entsprechend gedeutet werden.

Es scheint schließlich nicht gerechtfertigt, (a) das Zeigen in qualitativer Hinsicht mit der Sprache gleichzustellen und (b) das Zeigen auf der Grundlage eines gemeinsamen begrifflichen Hintergrunds mit dem Zeigen auf der Grundlage eines nur gemeinsamen Aufmerksamkeitsrahmens gleichzustellen. Tomasello (2002: 49) selbst stellt bei der Einordnung der Leistungen akkultrierter Affen fest, dass „auf eine Kultur zu reagieren und eine Kultur *de novo* zu erschaffen, [...] zwei verschiedene Dinge [sind]“, und spricht damit ungewollt ein auf seinen eigenen Ansatz übertragbares Problem an. Die Verwendung signalartiger Kontakte in Reaktion auf ein symbolisch konstruiertes Umfeld wird nämlich kaum entscheidende Hinweise zum Übergang von signalhaften zu symbolischen Kommunikationsstrukturen liefern.

## 5. Kommunikationsmittel im Kommunikationsprozess

Tomasellos kreativer Forschungsansatz hat sowohl im Bereich der Primatenforschung als auch hinsichtlich der Ontogenese der kindlichen Sozialkognition und des Spracherwerbs zweifellos interessante Ergebnisse generiert, die auch sozialtheoretisch von Nutzen sein können. Aufgrund einer kognitionstheoretischen Grundorientierung liefert aber selbst Tomasellos Angebot einer ‚Prototheorie der Handlungskoordination‘ in *Die Ursprünge der menschlichen Kommunikation* trotz angedeuteter Berücksichtigung von Sozialität durch Summation je individueller Kapazitäten lediglich eine Theorie der Kommunikationsmittel in ihrer Abhängigkeit von bestimmten kognitiven Vermögen. Im Hinblick auf den jeweiligen kognitiven Entwicklungsstand von Kindern bzw. die generelle kognitive Leistungsfähigkeit nichtmenschlicher Primaten generiert Tomasellos Forschung wertvolle Erkenntnisse, dennoch ist bereits das einseitige Bedingungsverhältnis von kognitiven Vermögen (Bedingung) und Kommunikationsmitteln (Konsequenz) problembehaftet: Die Zeigegeste wird nämlich als ursprünglichste Form humanspezifischer Kommunikation betrachtet, wenn sie von einem über die psychische Infrastruktur geteilter Intentionalität verfügenden Menschen verwendet wird. Da man Tomasello zufolge aufforderndes Zeigen aber auch bei nichtmenschlichen Primaten antreffen kann und die Zeigegeste Kindern motorisch zudem schon sehr viel früher verfügbar ist, muss unabhängig von der Nichtberücksichtigung des Prozessgeschehens, in dem die Kontaktmittel erst tatsächliche Qualitäten annehmen, die Frage erlaubt sein, ob das Zeigen überhaupt als Indikator des Umschlagpunktes zu humanspezifischen Kommunikationsmitteln taugt. Ursache dieses Problems mag erneut die Nichtberücksichtigung nebeneinander existierender Ebenen von Intentionalität sein. Keineswegs gesichert ist folglich Tomasellos evolutionäre Hypothese, nach der die ersten nur beim Menschen vorkommenden Formen der Kommunikation im Zeigen und im Gebärdenspiel lagen (Tomasello 2009: 13), zumal erst der synsemantische Einsatz der Kommunikationszeichen einen tatsächlichen Qualitätssprung markiert.

Schwierigkeiten in Tomasellos Ansatz entstehen in großem Maße dadurch, dass die schon für sich genommen qualitativ uneindeutig bestimmten Kommunikationsmittel (im Grunde aber die ihnen nicht ganz rechtmäßig zugeordneten kognitiven Vermögen) den Kommunikationsprozess (dem Tomasello sich kaum einmal näher zuwendet) einseitig fundieren sollen. Tatsächlich ist das Verhältnis von Kommunikationsmitteln und Kommunikationsprozessen aber ein Verhältnis gegenseitiger Fundierung, „da kein Kommunikationsprozeß ohne semantische Mittel auskommt, semantische Mittel ihrerseits aber wechselseitiger Verhaltenssteuerung erwachsen und deshalb nicht ohne Bezug zum Kommunikationsprozeß verstanden werden können.“ (Loenhoff 2003: 180) Die Nichtberücksichtigung dieses Sachverhalts führt in Tomasellos Ansatz nicht nur Missverständnisse hinsichtlich des Kommunikationsprozesses und der Rolle der Kommunikationsmittel innerhalb dieses Prozesses herbei, sie erschwert auch Aussagen zur Genese



der Kommunikationsmittel. Wird der Versuch unternommen, die Ganzheitlichkeit des Kommunikationsprozesses durch eher allgemeine Rückwirkungen der kulturellen Umwelt zu erfassen, obwohl Kommunikationsmittel wechselseitigen, semiotisch strukturierten Steuerungsprozessen (von der einfachen Verhaltensabstimmung bis zu den komplexesten Formen symbolisch strukturierter Handlungskoordination) erst erwachsen, sind Aussagen zur phylogenetischen und zur ontogenetischen Entstehung von Kommunikationsmitteln höchst anfällig. Dies gilt gleichermaßen für die Genese kognitiver Leistungen, und zwar selbst dann, wenn man von dem problembehafteten einseitigen Bedingungsverhältnis zwischen kognitiven Vermögen und Kontaktmitteln absieht.

Tomasellos Ansatz kann letztlich kaum ansprechende Antworten liefern, wenn es zu klären gilt, welche Rolle den identifizierten kognitiven Größen im Rahmen des Interaktionsprozesses tatsächlich zukommt, wie sie zum Zwecke der Handlungskoordination eingesetzt werden und wie sie selbst zum Gegenstand der Handlungskoordination werden können. All diese Fragen sind nur zu beantworten auf der Grundlage der in der Kognitionspsychologie vernachlässigten Frage nach dem tatsächlichen Ablauf der Handlungskoordination, der sowohl für Kinder verschiedener Altersstufen als auch für nichtmenschliche Primaten erst einmal zu beschreiben wäre, um nähere Hinweise zum Übergang von signalartigen zu (humanspezifischen) symbolischen Steuerungsprozessen zu gewinnen. In diesem Kontext könnten dann durchaus die von Tomasello zur Diskussion gestellten kognitiven Vermögen Berücksichtigung finden, zumal deren Genese ja nur im Kontext wechselseitiger Steuerungsprozesse sinnvoll nachzuvollziehen ist, ohne sich in zirkuläre Erklärungsmodelle flüchten zu müssen. Im Rahmen eines sozialpsychologisch fundierten (zunächst funktionalistisch bestimmten) Forschungsprogramms könnten die von Tomasello entschlüsselten psychischen Komponenten – die Fähigkeit(en), gemeinsame Intentionen, gemeinsame Aufmerksamkeit und einen (zunächst nichtbegrifflichen) gemeinsamen Hintergrund auszubilden – also fraglos zu Richtwerten werden, so f e r n man sich stets im Klaren ist, dass sie zwar Voraussetzung zum Beispiel des Einsatzes symbolischer Kontaktmittel sein könnten, jedoch keine eindeutige Zuordnung von Kommunikationsmitteln und kognitiven Komponenten vorliegt, dass außerdem allein die Identifikation eines speziesspezifischen oder in einem bestimmten Alter vorhandenen Vermögens weder verlässliche Aussagen über dessen (phylo- oder ontogenetische) Genese noch über den Prozess der Handlungskoordination herbeiführt.

## **6. Joint Attention: ein möglicher Anknüpfungspunkt?**

Bis hierhin hat sich eine funktionalistische Betrachtung der Qualität der Kontaktmittel selbst als äußerst hilfreich erwiesen, die aus Tomasellos mentalistischen Fundamenten resultierenden Einschränkungen und Fehlschlüsse aufzudecken. V o r s p r a c h l i c h e n Prozessen gemeinsamer Aufmerk-

samkeit könnte allein auf dieser Basis aber keine eigenständige, humanspezifische Qualität zugesprochen werden. Noch rein dyadisch ist der Kontakt zwischen einander drohenden Wölfen, bereits bei der gemeinsamen Jagd steuern die Tiere sich aber nicht nur gegenseitig, sondern auch mit Blick auf einen weiteren Steuerungsrichtpunkt, nämlich das Beutetier. Auch der auffordernd zeigende Affe deutet auf einen zusätzlichen Bezugspunkt, und selbst die tänzelnde Honigbiene gibt ihrer Artgenossin Anweisungen bezüglich einer Nahrungsquelle, die neben den Individuen selbst als weiterer Richtpunkt der Steuerung fungiert. Ganz ohne problematische Vorentscheidungen gerät auf der Grundlage einer funktionalistischen Betrachtungsweise also eine schon im Tierreich anzutreffende Form signalartiger Kontakte in den Blick, die wenigstens im Hinblick auf die Zahl der Steuerungsrichtpunkte über rein dyadische Steuerungen hinausgeht. Ablösbar von den Dingen und entstofflicht, zugleich losgelöst von den empraktischen Bezügen sind die Kontaktmittel hier fraglos nicht; dies gilt zugleich aber für vorsprachliche Akte gemeinsamer Aufmerksamkeit.

Allerdings soll die besondere Qualität von Prozessen gemeinsamer Aufmerksamkeit, deren Erforschung von Tomasello mitbegründet und entscheidend vorangetrieben wurde, hier nicht geleugnet werden. Im Sinne einer Orientierung an der Orientierung anderer (Knobloch 1997) gilt Joint Attention gemeinhin als Quellpunkt symbolischer Kommunikation. Sie ist die Grundlage einer Entwicklung von der Verwendung einfacher Signale hin zum Gebrauch konventionalisierter Symbole und zugleich die Grundlage des Prozesses der Dekontextualisierung als einer Entwicklung von der pragmatischen, auf der Kategorisierung situativer Kontexte beruhenden Verwendung sprachlicher Ausdrücke zu einer Kategorisierung der Ausdrücke innerhalb der Sprache selbst (Kathage 2008: 70; Feilke, Kappesst u.a. 2001: 10). Sprachliche Ausdrücke können gar als eine Art Verkörperung von Joint-Attention-Prozessen angesehen werden, da sie die Haltung von Hörer und Sprecher bezüglich eines dritten Gegenstandes widerspiegeln (Knobloch 1998: 252; Kathage 2008: 70).

Es ergibt sich an diesem Punkt also die Notwendigkeit, mentale Phänomene in den Blick zu nehmen, wenn der Übergang zur symbolischen Kommunikation näher aufgeschlüsselt werden soll. Bühler vollzieht mit dem Übergang zur humanspezifischen symbolischen Steuerung einen Perspektivenwechsel, den genauen onto- und phylogenetischen Vollzug dieses Umbruchs beschreibt er in seiner *Krise der Psychologie* aufgrund einer anderweitigen Zielsetzung aber nicht im Detail. Dass er keinen strikten Antimentalismus vertritt, verdeutlicht er, indem er die Erlebnispsychologie in seine Axiomatik aufnimmt. Zugleich bemerkt er jedoch, dass ein derart ausdruckstheoretisches Modell wie die Erlebnispsychologie nicht als geeignete Ausgangsbasis taugt, dass im Hinblick auf mentale Phänomene also stets deren Einbettung in Sozialität zu berücksichtigen ist. Demgemäß ist Tomasello die Erkenntnis anzurechnen, dass im Verlauf der von ihm sogenannten Neunmonatsrevolution individuelle Kompetenzen generiert werden, die für den Erwerb humanspezifischer – und damit auch funk-

tional einzigartiger – Kommunikationsformen zwingend erforderlich sind; werden jedoch Intentionen, das Verstehen von Intentionen oder die letztlich über das Verstehen anderer als kooperative Akteure definierte geteilte Intentionalität zu voraussetzungslosen Ausgangspunkten erklärt, bleiben bedeutende Aspekte der Interaktion unberücksichtigt und es kommt zu den oben behandelten Erklärungswidersprüchen. Jürgen Habermas (2012: 61–65) stellt daher die Notwendigkeit des „Dazwischentretens von Gesten“ heraus. Habermas adressiert die auch von Tomasello angesprochene triadische Beziehung, die durch Verknüpfung horizontaler Beziehungen zwischen den Interaktanten mit dem vertikalen Weltbezug entsteht und zunächst mit dem Einsatz von Gesten verbunden sei. Erst die Verschränkung von Wahrnehmungen und Blickrichtungen schaffe hier den objektivierenden Bezug zur Welt und erst der in diesem Rahmen stattfindende Abgleich reziprok übernommener Perspektiven löse die noch bei nichtmenschlichen Primaten gegebene egozentrische Wahrnehmung der Umgebung ab. Einer Fähigkeit zur gestenvermittelten Bezugnahme auf objektive Gegebenheiten, bei der die Interaktanten dieselben Ziele verfolgen, müsse im nächsten Schritt aber noch eine Konventionalisierung des Zeichengebrauchs, durch die das Zeichensubstrat erst zum Träger von Bedeutung erhoben werde, folgen. Symbolische Gehalte könnten also nur dann entstanden sein, wenn geteiltes Wissen regelmäßig mit jenen Lauten und Bewegungen assoziiert worden sei, die der bloßen Herstellung gemeinsamer Aufmerksamkeit zunächst als Katalysator dienen. Zwar veranschlagt auch Habermas sozialkognitive Voraussetzungen, er distanziiert sich aber entscheidend vom mentalistischen Ansatz Tomasellos, der die symbolischen Bedeutungen auf geteilte Intentionen und Wahrnehmungen zurückführe: „Der mentalistische Ausdruck ‚geteiltes Wissen‘ verdeckt die welterschließend-konstituierende Leistung der Symbolisierung, die wir von der sozialkognitiven unterscheiden müssen.“ (Habermas 2012: 63; Hervorhebungen im Original) Ein Vorrang mentaler Phänomene vor der Kommunikation verbiete sich, denn erst über das Dazwischentreten von Gesten lasse sich die für die Entstehung symbolischer Kommunikation erforderliche Verschränkung der interpersonalen Beziehung mit einer objektivierenden Einstellung zur Welt erklären. Nur im Prozess der Handlungscoordination kann gemeinsame Aufmerksamkeit (einergehend mit entsprechenden psychischen Schlüsselkomponenten) zuerst signalartig konstituiert werden, um als Steuerungsprozess der Genese symbolischer Akte als Fundament zu dienen.

Es mag schließlich durchaus sinnvoll sein, Joint Attention als eigenständige, bereits humanspezifische Steuerungsqualität auszuweisen. Zwar ist die von vielen Theoretikern hervorgehobene triadische Beziehung unabhängig von intentionalen Gesichtspunkten auch im Tierreich anzutreffen, doch scheinen in triadischen Prozessen der Handlungscoordination, die der kindlichen Einbringung in symbolisch strukturierte Steuerungen unmittelbar vorausgehen, bestimmte kognitive Vermögen in der Koordination von Wahrnehmungsaktivitäten verhaltensmäßig zum Ausdruck zu kommen, die

dem Tierreich verborgen bleiben und die somit einen bedeutenden Anteil am Übergang zu humanspezifischer Kommunikation haben könnten. Weil es nur symphysisch und empraktisch funktioniert und selbst die Aspekte der Arbitrarität und Konventionalisierung im Grunde auch von akkulturierten Affen erfüllt werden (Mollenhauer 2010), ist das einzelne Sprachzeichen als Spiegel einer nur dem Menschen vorbehaltenen gemeinsamen Aufmerksamkeit ungeeignet. Vorerst ist also von einer humanspezifischen Qualität einer synsemantischen Zeichenverwendung auszugehen, die aus auch im Tierreich anzutreffenden signalartigen (triadischen) Steuerungen hervorgeht, (vermutlich) wenn bestimmte kognitive Vermögen hinzutreten.

Nähere Aussagen zur Genese humanspezifischer Kommunikation und zur Genese der beteiligten kognitiven Ressourcen verlangen nach einer empirischen Forschung, die sich eben nicht auf den Nachweis kognitiver Kompetenzen beschränkt, sondern stattdessen schon in ihrer Anlage auf die Beschreibung des Prozesses der Handlungskoordination ausgerichtet ist. Nur eine Beschreibung der Koordinationsaktivitäten kann letztlich auch die Unterschiede zwischen (funktional) triadischen Steuerungsaktivitäten im Tierreich und Aktivitäten gemeinsamer Aufmerksamkeit beim Menschen entschlüsseln, während Befunde *allein* zu kognitiven Aspekten oder der Verwendung gewisser Steuerungsmittel in Bezug auf den Steuerungsprozess und die dort tatsächlich abgerufenen kognitiven Ressourcen kaum aussagekräftig sind. Verbunden mit dieser Erkenntnis ist das Plädoyer für (a) eine Forschung zur kommunikativen Entwicklung des Kindes sowie (b) eine Art ‚Vergleichende Interaktionsforschung‘. Die Erforschung der kommunikativen Entwicklung des Kindes rückt im Gegensatz zur weithin psychologisch (und nunmehr verstärkt auch neurowissenschaftlich) geprägten Spracherwerbsforschung nicht das einzelne Kind und seinen Erwerb sprachlicher Strukturen in einer Umwelt in den Mittelpunkt, sondern verfolgt einen ganzheitlichen Ansatz, der die Entwicklung des Kindes im Rahmen eines nicht weiter zerlegbaren multimodalen Interaktionsgeschehens erforscht. Nur in dieser sozialen Praxis kann auch die kognitive bzw. sozial-kognitive Entwicklung des Kindes vonstattengehen. Auch der im Wesentlichen an der kognitiven Entwicklung interessierte Psychologe kommt um eine ganzheitliche Berücksichtigung natürlicher Interaktionen demnach nicht herum. In vergleichenden Studien mit nichtmenschlichen Primaten könnte über die bloße Feststellung einer Beschränkung auf signalartige Kontakte hinaus erforscht werden, welche Koordinationsmuster wir mit unseren nächsten lebenden Verwandten teilen, was möglicherweise nähere Aufschlüsse zum Übergang zu humanspezifischen Interaktionsformen erlaubt.

## Anmerkungen

- 1 Die nachfolgenden Ausführungen skizzieren die zeichentheoretisch relevanten Ergebnisse meiner Dissertation, die unter dem Titel *Tomasellos Kooperationsmodell*.

*Michael Tomasellos Forschung im Kontext kommunikationstheoretischer Fragestellungen* (2015) im Verlag UVK erschienen ist.

## Literatur

- Bühler, Karl (1999). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. 3. Auflage. Stuttgart: Lucius & Lucius. [Zuerst erschienen 1934].
- Bühler, Karl (2000). *Die Krise der Psychologie*. Karl Bühler Werke 4. Weilerswist: Velbrück Wissenschaft. [Zuerst erschienen 1927].
- Feilke, Helmut, Klaus-Peter Kappes und Clemens Knobloch (2001). Grammatikalisierung, Spracherwerb und Schriftlichkeit – Zur Einführung ins Thema. In: dieselben (eds.). *Grammatikalisierung, Spracherwerb und Schriftlichkeit*. Tübingen: Niemeyer, 1–28.
- Habermas, Jürgen (2012). Die Lebenswelt als Raum symbolisch verkörperter Gründe. In: derselbe. *Nachmetaphysisches Denken II. Aufsätze und Repliken*. Frankfurt a.M.: Suhrkamp, 54–76.
- Kathage, Andrea (2008). *Zur kommunikativen Entwicklung des Kindes. Grundlagenstudie zur Entwicklung eines kommunikationswissenschaftlichen Ansatzes*. Aachen: Shaker.
- Kleist, Heinrich von (1984). Über die allmähliche Verfertigung der Gedanken beim Reden. In: Heinrich von Kleist. *Sämtliche Erzählungen und andere Prosa*. Stuttgart: Reclam, 340–346.
- Knobloch, Clemens (1997). Rezension von Boris M. Velichkovsky & Duane M. Rumbaugh (eds.), *Communicating Meaning: The Evolution and Development of Language*. *Sprache und Kognition* 16, 2, 127–132.
- Knobloch, Clemens (1998). Reference: Grammaticalizing Joint Attention. *Pragmatics and Cognition* 6, 1–2, 245–264.
- Loenhoff, Jens (2003). Kommunikationstheorie und Fundierungsrelationen im interpersonellen Kommunikationsprozeß. In: Helmut Richter und H. Walter Schmitz (eds.). *Kommunikation: ein Schlüsselbegriff der Humanwissenschaften*. Münster: Nodus, 179–191.
- Loenhoff, Jens und Rafael Mollenhauer (2016). Zwischen Kooperation und methodologischem Individualismus. Zu den kognitionstheoretischen Hintergründen von Tomasellos Kommunikationsbegriff. In: Gert Albert, Jens Greve und Rainer Schützeichel (eds.). *Kooperation, Sozialität und Kultur. Michael Tomasellos Arbeiten in der soziologischen Diskussion*. 3. Sonderband der *Zeitschrift für Theoretische Soziologie*, 102–127.
- Mead, George Herbert (1973). *Geist, Identität und Gesellschaft aus der Sicht des Sozialbehaviorismus*. Frankfurt a.M.: Suhrkamp.
- Mollenhauer, Rafael (2010). *Symbolverwendung bei Primaten? Eine Analyse der Ansätze von David Premack und Susan Savage-Rumbaugh*. Aachen: Shaker.
- Mollenhauer, Rafael (2015). *Tomasellos Kooperationsmodell. Michael Tomasellos Forschung im Kontext kommunikationstheoretischer Fragestellungen*. Konstanz: UVK.
- Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard: Harvard University Press.

- Tomasello, Michael (2009). *Die Ursprünge der menschlichen Kommunikation*. Aus dem Amerikanischen von Jürgen Schröder. Frankfurt a.M.: Suhrkamp.
- Ungeheuer, Gerold (2004). Die kybernetische Grundlage der Sprachtheorie von Karl Bühler. In: derselbe. *Sprache und Kommunikation*. 3., erweiterte und völlig neu eingerichtete Auflage. Herausgegeben und eingeleitet von Karin Kolb und H. Walter Schmitz. Münster: Nodus, 128–146.

*Dr. Rafael Mollenhauer*  
*Universität Duisburg-Essen*  
*Fakultät für Geisteswissenschaften*  
*Institut für Kommunikationswissenschaft*  
*Universitätsstr. 12*  
*D-45141 Essen*  
*E-Mail: rafael.mollenhauer@uni-due.de*