

Jörg R.J. Schirra,
Klaus Sachs-Hombach

Kontextbildung als anthropologischer Zweck von Bildkompetenz

Abstract

The anthropological function of pictures lies in conveying initial context building, i.e., the genesis of the capacity to focus on situations that are not currently present. In order to support that thesis, we first recapitulate the concept of communication with two ways of presentation enclosed therein which are central to language competence and picture competence. Thereby a path opens up to define perceptoid media in more detail, through which a transition to initial context building can be suggested. However, it becomes clear, too, that the ability to present oneself as a picture user has to be stabilized by further acts of communication. In the end, they also lead to the option of context building without a fallback on perceptoid media that we generally call visual imagination.

Die anthropologische Funktion von Bildern liegt darin, die initiale Kontextbildung zu vermitteln, d.h. die Fähigkeit, nicht aktuell anwesende Situationen zu vergegenwärtigen. Um diese These zu stützen, wird zunächst der Begriff der Kommunikation mit den darin enthaltenen beiden Arten des Zeigens rekapituliert, die für Sprach- und Bildkompetenz zentral sind. Damit eröffnet sich ein Weg, den Begriff wahrnehmungsnaher Medien näher zu bestimmen, über den ein Übergang zur initialen Kontextbildung vorgeschlagen werden kann. Allerdings ergibt sich, dass die Fähigkeit, sich als Bildbenutzer zu zeigen, durch weitere Kommunikationsakte stabilisiert werden muss. Sie führen letzt-

lich auch zu der Option von Kontextbildungen *ohne* Rückgriff auf wahrnehmungsnahe Medien, die wir gemeinhin als *visuelle Vorstellungskraft* bezeichnen.

1. Kontextbildung als anthropologisches Spezifikum

Die Philosophie- und Wissenschaftsgeschichte bietet viele Merkmale zur Charakterisierung dessen an, was den Menschen ausmacht. Neben den biologischen Eigenheiten, die in der vorliegenden Betrachtung weniger relevant sind, dreht sich die Diskussion vor allem darum, sich mittels propositionaler Sprache zu verständigen und rational (d.i. ›politisch‹) organisierte Gesellschaften zu bilden. Auch über ›Vernunft‹, ›Geist‹, ›Kultur‹, ›Zivilisation‹ oder ›Symbole‹ zu verfügen und der Welt mit ›Verständnis‹ zu begegnen, ihr und ihren Teilen ›Bedeutung‹ zuzuschreiben, wird häufig genannt. Schließlich gehört die Fähigkeit, Bilder zu verwenden, der dieser Text vor allem gewidmet ist, ebenfalls hierher.

In der philosophischen Betrachtung richtet sich der Blick nicht auf eine Eigenart allein der biologischen Spezies ›homo sapiens sapiens‹, auch wenn empirisch bislang ausschließlich Exemplare dieser Art als Beispiele in Frage zu kommen scheinen. Vielmehr wird das Interesse auf eine prinzipielle Komplexität des Verhaltensrepertoires fokussiert, die auch bei Wesen ganz anderer Herkunft (etwa extra-terrestrischen Ursprungs) auftreten mag. Es geht, genau genommen, um den *Begriff*,¹ den wir uns vernünftiger Weise von Wesen bilden sollten, die zu einem entsprechend hoch-komplexen Verhalten der Welt, ihrem eigenen Körper und anderen Wesen eben dieser ›menschlichen‹ Art gegenüber fähig sind.

Indem sich der Blick auf den Begriff richtet, kommen statt bestimmter uns interessierender *Gegenstände* nun *unsere Gewohnheiten*, solche Phänomene zu unterscheiden, – und damit wir selbst auf doppelte Weise – in den Fokus: Sind wir doch einerseits – als Menschen – zugleich Beobachter und Beobachtetes, während wir andererseits unsere eigenen Unterscheidungsgewohnheiten, die bei jedem solchen Beobachtungsvorgang unumgänglich sind, kritisch betrachten. Ein Aspekt des philosophischen Interesses an der Anthropologie besteht neben anderen darin, die Menge der vorgeschlagenen begrifflichen Kriterien kritisch zu sichten, die Beziehungen dazwischen, wenn möglich, zu systematisieren und idealerweise eine Äquivalenzklasse minimaler Bedingungen zu bestimmen.² Für H. Plessner, einen der Begründer der philosophischen Anthropologie, ist eine solche Minimalbedingung bekannt-

¹ Unter einem Begriff wird hier, wie in der Philosophie im Anschluss an Wittgenstein üblich, eine Unterscheidungsfähigkeit, die interindividuell kontrolliert wird, verstanden (vgl. etwa ROS 1999). Entsprechend spielen empirische Untersuchungen – also das praktische Verwenden gewisser Unterscheidungsgewohnheiten – bei einer kritischen Sichtung von Begriffen eine zwar komplexe aber untergeordnete Rolle.

² Dass es sich hierbei nur um einen verhältnismäßig wenig bedeutsamen Aspekt der philosophischen Anthropologie handelt, wird etwa in der Einleitung von LORENZ 1990 deutlich.

lich die »exzentrische Positionalität«: Jedes Wesen, das unter diesen Begriff gefasst werden kann, müsste (und sollte) folglich etwa auch vor einem unserer Gerichtshöfe denselben Status haben, der einem Mitglied unserer eigenen Spezies unter vergleichbaren Bedingungen zugestanden wird. Dabei hat Plessner die Fähigkeit eines Wesens im Auge, zu sich selbst (insbesondere seiner leiblichen wie psychischen Existenz) »Distanz aufbauen« zu können.

Ein verwandtes Kriterium des Mensch-Seins ist die Fähigkeit zur Kontextbildung, d.h. zur Vergegenwärtigung nicht aktuell anwesender Situationen.³ Dass ein Wesen in der Lage ist, sein Verhalten nicht nur an seiner gerade vorhandenen Umwelt (im Sinne Uexkülls) auszurichten, sondern beliebige weitere potentielle Verhaltenssituationen ebenfalls zu berücksichtigen, ist in vielen typisch menschlichen Tätigkeiten eine entscheidende Voraussetzung. Alles, was mit vergangenen oder zukünftigen Situationen zu tun hat, wie auch alles Fiktive, Hypothetische und Kontrafaktische, wäre ohne Kontextbildung unmöglich. Selbst der Umgang mit Abstraktem – wenn also von konkreter raumzeitlicher Situierung abgesehen wird – scheint uns nur zu gelingen, wenn wir uns einen entsprechenden *metaphorischen* »Raum« für die betrachteten Abstrakta vorstellen (vgl. etwa LAKOFF 1990).

Machen wir uns klar, dass es bislang keine Notwendigkeit dafür gibt, die Fähigkeit, das jeweilige Hier und Jetzt als alleinigen Bezugspunkt des Verhaltens abzulösen, bei anderen Wesen außer den Primaten anzunehmen; und selbst in dieser biologischen Ordnung zeigt sie sich in vollem Umfang nur bei der Spezies *homo sapiens*. Ansatzpunkte gibt es wohl im Spielen der Tiere, insbesondere bei Jungtieren, wird doch dabei die Abhängigkeit des Verhaltens von der tatsächlichen Verhaltenssituation⁴ durch eine Abhängigkeit von einer nur dargestellten bzw. vorgespielten anderen Situation zumindest teilweise ersetzt. Doch fehlt die Distanz zu der Spielsituation, die für menschliche Kontextbildung charakteristisch ist: Ins Spiel vertieft, werden die meisten der Verhaltensweisen, die in der tatsächlichen Situation adäquat sind, vollständig ersetzt durch Verhaltensweisen, die in der fiktiven Spielsituation angemessen wären. Wenn etwa ein Welpen spielerisch einem Ball als einem Beutetier hinterher jagt, ist das, was er da jagt, für ihn ganz und gar ein Beutetier, nämlich gänzlich integriert in die entsprechenden Wahrnehmungs- und Verhaltensschemata. Zugleich bleibt das Spielverhalten aber auch strikt an die aktuelle Situation gebunden: Der als Beutetier genommene Ball etwa wird ja ganz korrekt in seiner physischen Erscheinung gehandhabt, seine Form, sein Gewicht etc. werden – im Rahmen der Fähigkeiten, über die das spielende Wesen im Umgang mit echter Beute verfügt – richtig einbezogen. Schließlich

³ Der Zusammenhang mit Plessners Exzentrizitätsbegriff ergibt sich dadurch, dass jene Distanzierung einer Betrachtung der aktuellen Situation als einer anderen (von »außen«) entspricht und damit eine Kontextbildung voraussetzt.

⁴ Genauer muss es hier heißen: der tatsächlich vom spielenden Wesen wahrgenommenen Situation inklusive seiner aktuell vorliegenden Wünsche und Ziele.

bleibt das Spiel der Tiere in der Regel auf die Heranwachsenden beschränkt und verschwindet bei den adulten Tieren weitgehend.⁵

Machen wir uns weiterhin klar, dass die Fähigkeit zur Kontextbildung notwendigerweise eine nur *sozial* zu etablierende Fähigkeit ist: Von einem ›geistigen‹ Zugang eines Wesens zu einer nicht-aktuellen Situation im hier gemeinten Sinn kann nämlich nur dann sinnvoll die Rede sein, wenn es sich eben dieser Situation nicht nur ein einziges Mal, sondern wiederholt zuwenden kann. Dass es sich dabei tatsächlich um *dieselbe* Situation (mit denselben, nicht nur gleichartigen Gegenständen) handelt, kann aber wiederum nur sichergestellt werden, wenn ein anderes Individuum (zumindest im Prinzip) diese Kontextbildungen kontrollieren und dabei beurteilen kann, dass sie regelgerecht durchgeführt und also auf denselben Kontext bezogen sind. Daher auch ist eine Verhaltenssituation – eine ›Umwelt‹ – nur dann ein Kontext im hier gemeinten Sinn, wenn sie als (zumindest prinzipiell) interindividuell verfügbar betrachtet werden kann. Kontextbildung vollzieht sich also stets *kommunikativ*.

Sprache ist ein mächtiges Instrument der Kontextbildung, wenn nicht das mächtigste überhaupt. In dem anspruchsvollen Sinn von Sprache, der dabei üblicherweise unterstellt wird, beruht Sprachfähigkeit allerdings bereits selbst im wesentlichen auf der Fähigkeit zur Kontextbildung: Die Differenzierung in Nomination und Prädikation, die charakteristisch für Aussagen und daraus abgeleitete Sprechakte ist, zielt letztlich vor allem darauf ab, sich über Nicht-Gegenwärtiges zu verständigen. Insbesondere die Nomination, das sprachliche Identifizieren eines dem Gegenüber als bereits bekannt vorausgesetzten Gegenstandes, über den etwas (vermutlich) Neues mitgeteilt werden soll, ist eine Teilhandlung, die nur relativ zu einem vorgegebenen Kontext durchgeführt werden kann (TUGENDHAT 1976): Wer denn gemeint ist, wenn etwa »der Vater von Gesine« als nominatorischer Teil einer Aussage verwendet wird, hängt ganz davon ab, von welcher Situation eigentlich die Rede ist.⁶ Entsprechend sind unsere Sprachen voller Möglichkeiten zur Kontextbildung: Ausdrückliche Angaben von Ort und Zeit in Satzadverbien etwa, oder implizite Hinweise in Tempus (z.B. Futur) und Modus (z.B. Irrealis) des Verbs.

Kontextbildung spielt weiterhin eine zentrale Rolle für die Kompetenzen, mit individuierten persistenten Gegenständen umgehen und Ähnlichkeiten zwischen ihnen beurteilen zu können. Auch diese Gegenstände sind uns empirisch ja stets nur in ihrer momentanen Erscheinungsweise in der je aktuellen Verhaltenssituation gegeben. Der Aufbau der Fähigkeit, mit individuellen Gegenständen Situationen übergreifend umgehen zu können – kurz: Objektkonstitution – setzt die simultane Assoziation von Erscheinungsweisen

⁵ Ausnahmen finden sich interessanter Weise vor allem bei stark domestizierten Arten, die in besonders viele Interaktionen mit Menschen eingebunden sind (insbes. Hunde, Katzen).

⁶ Es könnte sich u.a. um eine Situation aus dem Komplex der Situationen handeln, die Uwe Johnson in seinem Roman *Jahrestage* (1970–1983) geschaffen hat: In dem Fall wäre etwa folgende Aussage wahr, die im Bezug auf andere Kontexte durchaus unwahr sein könnte: »Der Vater von Gesine ist Schreiner«.

aus verschiedenen Situationen zu einem die gegenwärtige Merk- und Wirkwelt transzendierenden ausdauernden Individuum voraus.

Andererseits macht der anspruchsvolle Ähnlichkeitsbegriff, den wir mit dem Begriff darstellender Bilder verbinden, Gebrauch von einem Vergleich zweier Situationen: die tatsächliche aktuelle Situation mit einem flachen, mit Pigmenten bedeckten Gegenstand (dem Bildträger) und den in dieser Situation angemessenen Reaktionen einerseits; und eine veränderte Situation, die statt des Bildträgers das mit ihm Dargestellte – nicht zuletzt individuelle Gegenstände – umfasst und die dazu passenden Reaktionen andererseits. Würde sich das Verhalten nur an einer dieser Situationen ausrichten statt an der Beziehung zwischen beiden, so könnte nicht vom Verhalten einem *Bild* gegenüber gesprochen werden: Im ersten Fall wird der Bildträger als ein Gegenstand ohne Bildcharakter betrachtet, im zweiten handelt es sich um eine einfache, nicht durchschaute Täuschung. Tatsächlich wirkt also der Gebrauch eines darstellenden Bildes als eine Kontextbildung, die in der einen Situation für die Bildnutzer die andere evoziert.

Damit ist der Rahmen für das in diesem Text behandelte Thema gesteckt: Es ergibt sich nämlich die philosophisch-anthropologische Kernfrage, wie es zu der zentralen Fähigkeit der Kontextbildung gekommen ist. Dabei ist diese Frage nicht im Sinn einer (prä)historischen, empirisch zu belegenden Chronologie zu verstehen. Vielmehr geht es um einen begriffsgenetischen Zusammenhang: Mit welcher rationalen Rekonstruktion können *wir uns* die Situation der *initialen* Kontextbildung begreiflich machen? Die dabei aufgezeigte logische Struktur der initialen Kontextbildung mag dann als Wegweiser für die Interpretation empirischer Funde dienen.

Die anthropologische Funktion des Bildes, so die These dieser Arbeit, liegt darin, die initiale Kontextbildung zu vermitteln. Zu diesem Zweck wird zunächst der Begriff der Kommunikation mit den darin enthaltenen beiden Arten des Zeigens rekapituliert, die sowohl für Sprache wie Bildkompetenz zentral sind (2). Damit eröffnet sich ein Weg, den Begriff wahrnehmungsnaher Medien näher zu bestimmen, über den ein Übergang zur initialen Kontextbildung vorgeschlagen werden kann (3). Allerdings ergibt sich, dass die Fähigkeit, sich als Bildbenutzer zu zeigen, durch weitere Kommunikationsakte stabilisiert werden muß. Sie führen letztlich auch zu der Option von Kontextbildungen *ohne* Rückgriff auf wahrnehmungsnahe Medien, die wir gemeinhin als *visuelle Vorstellungskraft* bezeichnen (4).

2. Kommunikation und der Doppelsinn des Zeigens

Da Kontextbildung als ein Akt der Kommunikation betrachtet werden soll, lohnt es sich, zunächst begrifflich zu rekapitulieren, was genau mit ›Kommunikation‹ gemeint ist.

nikation« gemeint ist.⁷ Unter Kommunikationshandlungen verstehen wir spezielle Formen von interaktiven Handlungen. Letztere wiederum sind zusammengesetzte Handlungen mit mindestens zwei Handelnden: Als Aktivitätsträger tritt bei einigen der Teilhandlungen der eine, bei anderen der andere Handelnde auf.⁸ Zudem enthalten diese Teilhandlungen Komponenten, die in einem sehr weiten Sinn jeweils auf den anderen beteiligten Handelnden ausgerichtet sind: Es handelt sich daher um »Beteiligungshandlungen«. Dazu genügt es, dass die Teilhandlung vom Anderen wahrgenommen werden kann.⁹

Sinnvollerweise sollte zwischen *kommunikativen*, *direkten* und *sachbezüglichen* Interaktionen unterschieden werden. Die einfachste Form der interaktiven Handlung ist die direkte Interaktion: Ihre Teilhandlungen zielen unmittelbar auf den Körper des anderen Handelnden. Typische Beispiele sind Handlungen, mit denen Raubtiere ihre Beute greifen (sowie die Reaktionen, mit denen der drohende Körperkontakt zu vermeiden versucht wird): die Würgehandlung einer Boa, der Zugriff eines Seeadlers, der Biss einer Marienkäferlarve. Bei kommunikativen Interaktionen liegen die Dinge sehr viel komplizierter: Hier geht es darum, dass der eine Interagierende die anderen Interagierenden durch seine (Teil)handlungen mit einem Sachverhalt vertraut machen möchte, kurz: ihnen etwas *zeigt*. Genau betrachtet treten allerdings in der kommunikativen Form von Interaktion sogar zwei Aktualisierungen von »Zeigen« auf.

2.1 »Zeigen« und »sich zeigen«

Um diese Doppeldeutigkeit des Zeigens bei kommunikativen Interaktionen zu klären, hilft ein Blick auf die Unterschiede zwischen kommunikativen und sachbezüglichen Interaktionen:

Auch bei letzteren geht es darum, dass der eine Interagierende den anderen Interagierenden durch seine (Teil)handlung mit einem Sachverhalt in der Welt vertraut macht. Allerdings geschieht dieses Bekanntmachen hier ohne Absicht, gewissermaßen hinter seinem Rücken: Wenn etwa ein Jungtier wahrnimmt, dass seine Mutter an einer bestimmten Stelle trinkt, wird es dabei zugleich mit dem Sachverhalt vertraut, dass es an jener Stelle etwas zu trinken gibt.¹⁰ Oder wenn jemand einen anderen beim unvermittelten Öffnen

⁷ Im Weiteren folgen wir im Wesentlichen den Darstellungen in ROS 1979 (insbes. Kap. 6-9) und ROS 2005 (insbes. Abschn. VI.2.1 und VI.3.3).

⁸ Der hier zugrunde gelegte weite Handlungsbegriff setzt voraus, dass als Aktivitätsträger das gesamte Wesen, nicht nur eines seiner Subsysteme (etwa ein Reflexbogen) für das Verhalten verantwortlich gemacht werden kann.

⁹ Dieses Wahrnehmen der ersten Teilhandlung kann – als zweite Teilhandlung – bereits eine Interaktion vervollständigen.

¹⁰ Es ist zu beachten, dass alle sachbezüglichen Interaktionen immer auch direkte Interaktionen sind, oder besser: direkt-interaktive Komponenten haben, die meist aus Wahrnehmungsakten bestehen.

einer Truhe sieht, wird er dabei zugleich mit dem Sachverhalt vertraut gemacht, dass diese Truhe unversperrt ist.¹¹

Betrachten wir hingegen ein Verhalten, durch das ein Wesen einem Artgenossen die Anwesenheit eines ihnen gefährlichen Raubtiers signalisiert: Solche Verhaltensweisen sind etwa von Affen gut dokumentiert. Ein entsprechender Warnruf soll den Adressaten nicht in die Lage bringen, die Anwesenheit des Feindes nun selbst zu bemerken; vielmehr soll (im Sinn eines evolutionären Zwecks) bei ihm eine passende Flucht- oder Schutzreaktion gerade ohne das eigene Wahrnehmen des Sachverhalts ausgelöst werden. Das gelingt, indem sich der ›Sender‹ *darstellt als einer, der das Raubtier wahrgenommen hat*: Er zeigt dessen Anwesenheit, indem er sich selbst als einen zeigt, der den Feind wahrgenommen hat. Eben dies ist die für kommunikative Interaktionen charakteristische Dopplung des Zeigens. Analoges gilt für den oben erwähnten Beispielfall, wenn der Truhen-Öffner tatsächlich die Absicht hat, durch sein Handeln die Andere wissen zu lassen, dass diese Truhe unverschlossen ist. Dann öffnet er nämlich nicht einfach nur die Truhe, sondern stellt sich explizit der anderen gegenüber dar als einer, der diese unverschlossene Truhe öffnet. Aus diesem Grunde gilt dann eine solche Interaktion auch als kommunikativ.¹² Offensichtlich muß er dabei darauf achten, dass die andere diese Selbstdarstellung auch wahrnimmt, so dass durchaus damit zu rechnen ist, dass die eigentliche Ausführung des Öffnens sich vom Fall des Truhe-Öffnens als einer lediglich sachbezüglichen Interaktion erkennbar unterscheidet.

Kommunikation findet also nur dort statt, wo zwei Zeigehandlungen ineinander verschachtelt vorliegen. Die beiden Komponenten ›jemand zeigt (jemandem) etwas‹ und ›jemand zeigt sich (jemandem) als etwas‹, werden auch als Zeigen im Sinne von *Indikation* bzw. *Ostentation* bezeichnet.¹³ Die Ostentation ist dabei auf den Sender und seine Körperlichkeit bezogen: Er zeigt sich dem andern als ein bestimmter, der (im weiten Sinn) Schlüsse über sein Wissen und Wollen auslöst. Selbstpräsentation ist mithin eine notwendige Teilhandlung des Kommunizierens. Zugleich wird über diese Selbstpräsentation eine indirekte Zeigehandlung im Sinne von Indikation vollzogen, die auf etwas ganz anderes gerichtet ist. Im Gegensatz zu sachbezüglichen Interaktionen kann dabei das Ziel der Indikation auch außerhalb dessen liegen, was der ›Empfänger‹ selber wahrnehmen kann. Im oben erwähnten Beispiel der Affenwarnrufe ist das ganz evident.

¹¹ Dabei gehen wir hier davon aus, dass derjenige, der die Truhe öffnet, das nicht bereits mit der Absicht tut, den anderen sehen zu lassen, dass er sie öffnet, um ihm so mitzuteilen, dass sie offen ist.

¹² Für diese Einordnung spielt es keine Rolle, ob der andere erkennt, dass diese kommunikative Absicht besteht. Diese Bedingung kommt erst hinzu, wenn statt einzelner kommunikativer Interaktionen ganze Ketten davon mit wechselnden Rollen betrachtet werden.

¹³ Leider ist die Zuordnung von Indikation zu ›etwas zeigen‹ und Ostentation zu ›sich zeigen als‹ terminologisch derzeit noch nicht allgemeingültig fixiert. Die Wahl der umgekehrten Zuordnung wäre im Prinzip durchaus ebenfalls plausibel.

2.2 Zeichensprachliche Handlungen und propositionale Sprache

Auch kommunikative Interaktionen können weiter nach Komplexität unterschieden werden: Abgesehen von einfachen Ausdrucksaktivitäten, die, wie im Beispiel des Affen-Warnrufs, als Signale verwendet werden, sind hierbei vor allem die zeichensprachlichen Handlungen relevant, bei denen die Kommunizierenden selbst um den Sinn ihres kommunikativen Tuns wissen.¹⁴

Der Übergang von einfachen kommunikativen Interaktionen zu zeichensprachlichen Handlungen ist dadurch charakterisiert, dass das Vorführen von Verhaltensweisen, wie es für den Vollzug kommunikativer Handlungen wesentlich ist, »verinnerlicht« wird (vgl. MEAD 1934): Präziser gesagt geht es darum, dass sie »sich hier aus den kommunikativen Zusammenhängen, in die sie [...] normalerweise eingebettet sind, herauslösen, und zur Fähigkeit einer allein für sich zu vollziehenden leiblichen Selbstvergegenwärtigung entwickeln« können (ROS 2005: 591). Während ein in kommunikativer Absicht vorgeführtes Verhalten unter anderem so modifiziert wird, dass seine körperlichen Aspekte besonders deutlich und für den Empfänger gut erkennbar artikuliert werden, führt die Reflexivierung des Zeigens eines Verhaltens zu einer Reduktion seiner von außen klar erkennbaren Aspekte: Statt deutlich sichtbarer körperlicher Veränderungen (etwa das bei Wölfen als Geste aus dem Anfangselement einer Bissattacke entwickelte hervorgehobene Flatschen der Zähne) genügt nun eine sehr schwache Aktivierung der entsprechenden Nervenbahnen, eine kaum merkbare Änderung des zugehörigen Muskeltonus.¹⁵ Zudem wird die Aktivität des Vorführens des einen Kommunikationspartners – seiner Ostentation – systematisch mit den zugehörigen Reaktionen des anderen Kommunikationspartners hinsichtlich der Indikation in Beziehung gesetzt: Begrifflich kann so die Rolle des Gegenübers mit der Rolle des ersten Handelnden in einen stabilen Zusammenhang gebracht werden. Damit sind Individuen, die in bewusste Kommunikation eintreten können, stets als dyadisch organisiert zu verstehen. Erst auf diese Weise werden kommunikative Handlungen als auf Regeln bezogene Handlungen (im Gegensatz zu Gewohnheiten) – und damit zeichensprachliche Handlungen – überhaupt möglich.¹⁶

Bestimmte Zeichenhandlungen ermöglichen es den Beteiligten, sich einem anderen gegenüber absichtsvoll darzustellen nicht nur hinsichtlich

¹⁴ Ros spricht hier von (im weiten Sinne) sprachlichen Handlungen. – Unter Signalen werden jene Handlungen verstanden, die automatisch situationsabhängig eingesetzt werden, weil sie – angeborener oder erworbener Maßen – zweckmäßig für eine Kommunikation wirken. Es sind kommunikativ eingesetzten Anzeichen. Demgegenüber setzen die eigentlichen Zeichenhandlungen (Ros: «sprachliche Handlungen») eine Intention zur Kommunikation voraus, sind also nicht bloß zweckmäßig, sondern werden spezifisch zu dem Zweck der Kommunikation eingesetzt.

¹⁵ Sie führen zwar kaum zu einer von außen noch unterscheidbaren Aktivität. Doch bleiben diese körperlichen Veränderungen innerlich wirksam über das propriozeptorisch vermittelte Körperbewusstsein, das ein wichtiges Element der Steuerung jedes körperlichen Verhaltens darstellt, das komplexer als einfache Reiz- Reaktionsschemata ist.

¹⁶ Es sei nochmals betont, dass es sich hierbei um eine begriffsgenetische, nicht um eine empirische Abfolge handelt; vgl. auch LORENZ 1990: 117 f.

eines jeweils *gegenwärtigen* Sachverhalts, sondern hinsichtlich eines *beliebigen* Sachverhalts oder besser: eines Sachverhalts relativ zu einem *beliebigen Kontext*. Eine solche Zeichenhandlung enthält eine *Proposition* – eine Teilhandlung im Gesamtgefüge der kommunikativen Selbstdarstellung, mit der jener Sachverhalt zum Ausdruck gebracht wird, dem gegenüber sich der eine Zeichenverwender den anderen präsentiert. Die Proposition vermittelt den Sachbezug der Gesamthandlung, während Interaktions- und Selbstbezug über andere Teile vermittelt werden.¹⁷ Diejenigen sprachlichen Zeichen, mit denen sich die Logik bereits seit langer Zeit auseinandergesetzt hat, Aussagen nämlich, gelten als die prototypische Form solcher propositionalen Zeichenverwendungen.

Eine Proposition besteht aus mehreren Teilhandlungen, die (i) verschiedene Funktionen erfüllen, (ii) auf den Kommunikationspartner bezogen sind und (iii) nicht unabhängig voneinander verwendet werden können: Mit der *Prädikation* versucht ein Sprecher verständlich zu machen, von welchen Fähigkeiten des Unterscheidens beziehungsweise Einordnens von Phänomenen – von welchen Begriffen also – er im Zusammenhang mit der aktuellen Gesamtzeichenhandlung Gebrauch machen möchte. Das ist der eigentlich informative Teil der Äußerung, der für die Kommunikationspartner neu sein sollte. Mit der (oder den) *Nomination(en)* versucht er verständlich zu machen, auf welchen einzelnen Gegenstand (oder Gegenstände) er sich dabei beziehen möchte: Diese Gegenstände sollten den Kommunikationspartnern bereits bekannt sein – als Teil des gemeinsamen ›Diskursuniversums‹.¹⁸ Aussagen bedürfen daher notwendiger Weise eines Aktes der Kontextbildung, da die Nominationen nur erfolgreich durchgeführt werden können, wenn allen Beteiligten klar ist, aus welcher Situation die gemeinten Gegenstände zur Identifikation ausgewählt werden (vgl. TUGENDHAT 1976: insb. §§ 21ff.).

Jemand, der einen Sachverhalt behauptet, stellt sich – sich selbst und anderen gegenüber – als einer dar, der die Wahrheit der Behauptung (rational) gegen Zweifel zu verteidigen bereit ist. Dabei kann die Überprüfung der Geltung auf zweierlei Wegen erreicht werden: Im günstigsten Fall ist die Proposition auf die Äußerungssituation bezogen (d.h. die Nominationen können einfach aufgelöst werden) und die Prädikation bringt eine Unterscheidungsgewohnheit ins Spiel, die direkt über Wahrnehmungskompetenzen (empirisch) entschieden werden kann (Kohärenztheorie der Wahrheit). Andernfalls ist zu unterscheiden, ob es möglich ist, den gemeinten Kontext für eine empirische Überprüfung körperlich zu erreichen, oder nicht. Im letzten Fall – wie in dem Fall, dass die Unterscheidungsgewohnheiten nicht ohne weiteres sinnlich eingelöst werden können – bleibt nur der zweite Weg, die neue Information mit dem bereits über jenen Kontext bekannten Wissen theoretisch-

¹⁷ Propositionen sind ›ungesättigte‹ Teilhandlungen (im Sinne Freges): Sie treten aus begrifflichen Gründen stets mit weiteren Teilhandlungen auf, die sie zu einer echten Zeichenhandlung ergänzen. Traditionell werden als Hilfsmittel, um sie dennoch in ›Reinform‹ darstellen zu können, ›dass‹-Phrasen verwendet, etwa ›dass dieser Satz aus 4 Wörtern besteht‹.

¹⁸ Das jeweilige Diskursuniversum ist ein von allen Gesprächspartnern geteilter und gerade gemeinsam fokussierter Kontext.

abstrakt zu vergleichen: Führt der Vergleich nicht zu logischen Inkonsistenzen, kann die behauptete Proposition als verteidigt gelten (konsistenztheoretischer Wahrheitsbegriff).

2.3 Das doppelte Problem der Sprachbegründung

Damit kehren wir zur Ausgangsthematik zurück: Ein großes anthropologisches Thema ist das Problem der *Sprachbegründung*. Der engen Kopplung von propositionaler Sprache und Kontextbildung wegen steht das Auftreten von Sprachkompetenz auch notwendig in unmittelbarem begrifflichem Zusammenhang mit der initialen Kontextbildung – dem Erwerb des Vermögens, einen nicht-gegenwärtigen Kontext der aktuellen Situation interindividuell entgegenzustellen. Hinweisende Definitionen spielten in den älteren, bewusstseinstheoretischen¹⁹ Ansätzen eine zentrale Rolle: Eine Gemeinschaft vorsprachlicher Individuen tauft einen gemeinsam wahrgenommenen Gegenstand mit einem weiterhin intersubjektiv verwendbaren Namen. Allerdings funktioniert Deixis unmittelbar nur in der jeweils aktuellen Situation.²⁰ Zudem bleibt ohne zusätzliche (sprachliche) Erläuterung unklar, was genau mit einer hinweisenden Geste gemeint ist. In der Tat zeigt sich die hier artikuliert Frage in zweierlei Gestalt in Kognitionswissenschaft und Philosophie: das Referenzproblem, das auch als Problem des *symbol grounding* bezeichnet wird, und das Problem der Etablierung interindividueller Bedeutung (*triangulatio ex nihilo*).

Das Referenzproblem, wie es insbesondere in den Kognitionswissenschaften diskutiert wird, besteht darin, näher zu bestimmen, wie die Beziehung zwischen sprachlichen Sachbezügen und den damit assoziierten Sachverhalten (bzw. der »Welt«) hergestellt werden kann. Vom bewusstseinstheoretischen Standpunkt, der den klassischen kognitionswissenschaftlichen Ansätzen in der Regel zugrunde liegt, werden sprachliche Strukturen durch mentale Symbole mit Bedeutung versehen, die ihrerseits als Elemente biologisch fundierter Informationsverarbeitungsprozesse verstanden werden. Doch wie können diese mentalen Symbole in der Welt »grundiert« werden?²¹ Das Problem des »symbol grounding« ist von den Grundlagen der bewusstseinstheoretischen Kognitionswissenschaften aus nicht zu lösen (vgl. WITTGENSTEIN 1922). Eine Alternative scheinen hier die auf phänomenologische Betrachtungen zurückgehenden Ansätze des »Enaktivismus« zu bieten, die das Referenzproblem als eine Frage des »Embodiments« betrachten (siehe etwa VARELA/THOMPSON/ROSCHE 1991; in aktuellerer Form: WHEELER 2005): Mentale Phänomene werden begriffen als wesentlich bestimmt durch ihre Einbettung in den

¹⁹ Die bewusstseinstheoretische Position versteht Begriffe als rein mentale Entitäten und völlig unabhängig von der Fähigkeit zu sprechen, so dass also die sprachliche Artikulation von Begriffen etwas hinsichtlich ihrer Funktion ganz Sekundäres bleibt und der Gebrauch von Begriffen (das Denken) auch ganz ohne Sprachvermögen möglich zu sein scheint.

²⁰ Die erweiterten Formen indirekter Deixis (Zeigen auf Stellen als metaphorischen Stellvertretern nicht-anwesender Dinge) sind begrifflich erst nach Einführung propositionaler Sprache möglich.

²¹ Searles Gedankenexperiment des »chinese room« hat dieses Problem prägnant veranschaulicht.

Leib und dessen Eingebundenheit in konkrete Situationen. Die mentalen Symbole erhalten ihre Bedeutung in einer wechselseitigen Anpassungsleistung zwischen Organismus und Umwelt. Damit das bewusstseinstheoretische Problem des Solipsismus vermieden wird, muß dabei allerdings die prinzipiell intersubjektive Natur der Bedeutungsphänomene berücksichtigt werden.

Donald Davidsons Versuch der Erläuterung, wie ein noch nicht ›bedeutungsinfiziertes‹ Kind in der Entwicklung zunehmend komplexerer Interaktionen mit bereits ›bedeutungsinfizierten‹ (erwachsenen) Gruppenmitgliedern den Umgang mit Bedeutung erwerben kann, scheint hier eine Lösung anzubieten (DAVIDSON 2001). Das bedeutungsgeladene Antwortverhalten der ›Erwachsenen‹ wird, grob gesagt, durch Training beim Kind mit dem eigenen Verhalten hinsichtlich eines Gegenstands in seiner Umwelt verbunden und internalisiert, wobei diese internalisierten Feedbacks die Basis der Bedeutung des Gegenstands (Zeichenträger) für das Kind werden. Eine solche ›semiotische Triangulation‹ setzt allerdings immer schon ein Umfeld mit Individuen voraus, die Sprache im anspruchsvollen Sinn bereits gemeistert haben. Wie aber soll man sich eine ›triangulatio ex nihilo‹ vorstellen, wie sie ja für die Individuen, die überhaupt erstmals die Kompetenz für ›bedeutungsvolles Sprechen‹ erwerben, benötigt würde?

Ebenfalls unklar bleibt, wie die semiotische Triangulation die Kompetenz der Kontextbildung begründen soll: Wie schon bei den hinweisenden Definitionen bleiben auch hier die bedeutungstiftenden Verhaltensweisen ja strikt an die je aktuelle Situation gebunden. Der Versuch Matthias Vogels einer semiotischen Triangulation für eine gegenüber propositionaler Rede elementarere Form des Kommunizierens (bzw. in seinen eigenen Worten: für nicht- sprachliche Gedanken) greift in dieser Hinsicht noch zu kurz, da auch die betrachteten prä-musikalischen oder prä-graphischen Ausdrucksformen einerseits noch vollkommen an den jeweiligen Äußerungskontext gebunden bleiben und andererseits keine Bedeutung außer ihrer selbst erwerben (VOGEL 2001). Der Ansatz weist aber in die von uns im Weiteren verfolgte Richtung, da er verständlich machen kann, wie der Begriff einer Kompetenz zu zwar kontextgebundenem, aber syntaktisch strukturiertem, gemeinsamem Verhalten mit »ästhetischen« Qualitäten eingeführt werden kann, das eine Grundlage für die Produktionsseite von im weiten Sinne bilderzeugenden Tätigkeiten als Basis für deren sozial koordinierte Rezeption liefert.

3. Die Rolle der Bildkompetenz

Unsere These lautet: Nicht-sprachliche ›ästhetische‹ Medien bilden in der Tat eine unumgängliche Komponente bei der initialen Kontextbildung und damit der begriffsgenetischen Begründung von propositionaler Sprachkompetenz. Um diesen Zusammenhang ausarbeiten zu können, ist es hilfreich, sich die spezifische Funktion von Bildern vor Augen zu führen: Ihre Verwendung zu

Kontextbildungen, die den Verwendern eine *gemeinsame empirische Vergegenwärtigung* abwesender Situationen ermöglicht.

3.1 Bilder als Mittel der empirischen Vergegenwärtigung

Obwohl Bilder bereits seit mehreren zehntausend Jahren im Leben der Menschen eine Rolle spielen, ist ihre Funktion auch heute keineswegs so klar, dass von einem breiten Konsens in der Bildphilosophie gesprochen werden könnte. Wir gehen im Folgenden von ihrer kommunikativen Funktion als grundlegend aus: ›*Wer sich oder einem anderen einen Gegenstand als ein Bild präsentiert vollzieht damit eine Kommunikationshandlung*«. Mit dieser Formulierung wird darauf hingewiesen, dass Gegenstände nicht einfach durch eigene physische Attribute als Bilder zu klassifizieren sind: Es ist ihre Rolle in einer spezifischen Verwendungssituation, die einen Gegenstand erst zu einem Bild werden lässt. Diese Verwendungssituation ist eine kommunikative Interaktion, die Rolle des Gegenstandes darin die des Zeichens.

Die obige Formulierung verweist zudem darauf, dass eine solche Interaktion *zwei* beteiligte dyadische Handlungsträger aufweist, die auch als zwei verschiedene Rollen von demselben Individuum (gleichzeitig oder auch zu unterschiedlichen Zeiten) eingenommen werden können.²² Zudem besteht eine zentrale Teilhandlung bei der Bildverwendung (wie bei jeder Kommunikation) notwendig in einer Ostentation, über die eine Indikation – der Verweis auf das mit dem Bild Bezeichnete – erst vollzogen werden kann. Die spezifische Beziehung zwischen Bezeichnetem und Bildträger hat diesen wichtigen Aspekt der Selbstdarstellung der Bildnutzer oft verdeckt. Dann scheint das Bild selbst – und das kann nur heißen: der Bildträger – etwas zu zeigen: eine stark verkürzte Redeweise, hinter der sich nur zu leicht die absurde Annahme vom Bildträger als einem eigenständigen Handlungsträger verbirgt.

Als ikonische Zeichen besteht bei Bildern, so heißt es etwa nach Peirce, eine perzeptuelle Ähnlichkeitsbeziehung zwischen dem Bildträger und dem mithilfe des Bildes Dargestellten. Um diese Ähnlichkeit in die handlungstheoretische Konzeption des Kommunikationsbegriffs einbeziehen zu können, bevorzugen wir den Ausdruck »wahrnehmungsnahes Zeichen« (SACHS-HOMBACH 2003): Wir gehen zu dessen näherer Bestimmung aus von einer schlichten Verwechslungsreaktion, die ein Wesen einem Teil seiner Umgebung gegenüber zeigt, ohne dabei die Verwechslung selbst noch zu bemerken – es erliegt einer Täuschung.²³ Der bei Bildverwendung vorliegende Zu-

²² Ein bildhaftes Zeichen kann demnach zur Kommunikation (i) mit sich selbst zur gleichen Zeit, (ii) mit sich selbst zeitversetzt oder (iii) mit einem anderen (gleichzeitig oder zeitversetzt) erfolgen. Die zeitversetzten Kommunikationen können dabei immer verstanden werden als doppelte, zu verschiedenen Zeiten stattfindende Teilkommunikationen mit jeweils vorgestelltem (also als gleichzeitig anwesend begriffenem) Kommunikationspartner. Im Folgenden werden wir von den zeitversetzten Verwendungsweisen absehen und uns auf die gleichzeitige Anwesenheit zweier Beteiligter konzentrieren.

²³ Der weit verbreitete Einsatz von Attrappen in der Ethologie beruht gerade auf solchen Verwechslungsreaktionen.

stand mit einer erkannten Ähnlichkeit ergibt sich durch die bewusste Assoziation der Verwechslungsreaktion (und der ihr adäquaten Verhaltenssituation) mit der tatsächlich die Täuschung auslösenden Situation (und den dazu adäquaten Verhaltensweisen). Einen Fall von Ähnlichkeit zu erkennen heißt also, die tatsächliche, in gewisser Hinsicht täuschende aktuelle Situation, die den »Ähnlichkeitsträger« enthält, mit einer anderen Situation in Verbindung zu bringen, in der die Täuschung keine Täuschung und die verwechselten Gegenstände tatsächlich anwesend wären.²⁴ Die spontanen Verhaltensweisen im Falle der Verwechslung bezeichnen wir als den *dezeptiven* Modus, in dem sich der Verwechselnde befindet, den Zustand des Erkennens von Ähnlichkeit hingegen als den *immersiven* Modus. Letzterer stellt das definierende Merkmal wahrnehmungsnaher Zeichen dar (SCHIRRA/SCHOLZ 1998). Dabei wird der immersive Modus über den dezeptiven Modus bestimmt: Das Erkennen einer Ähnlichkeit verweist auf die zugehörige, begrifflich »eingebettete« Verhaltensdisposition, die im Fall der Verwechslung spontan aktiviert wird. Sie bestimmt, wozu der betrachtete Teil der aktuellen Verhaltenssituation eigentlich ähnlich ist. Als erkannte Täuschung führt sie aber nicht – oder jedenfalls kaum noch – zu von »außen« wahrnehmbaren Reaktionen.²⁵

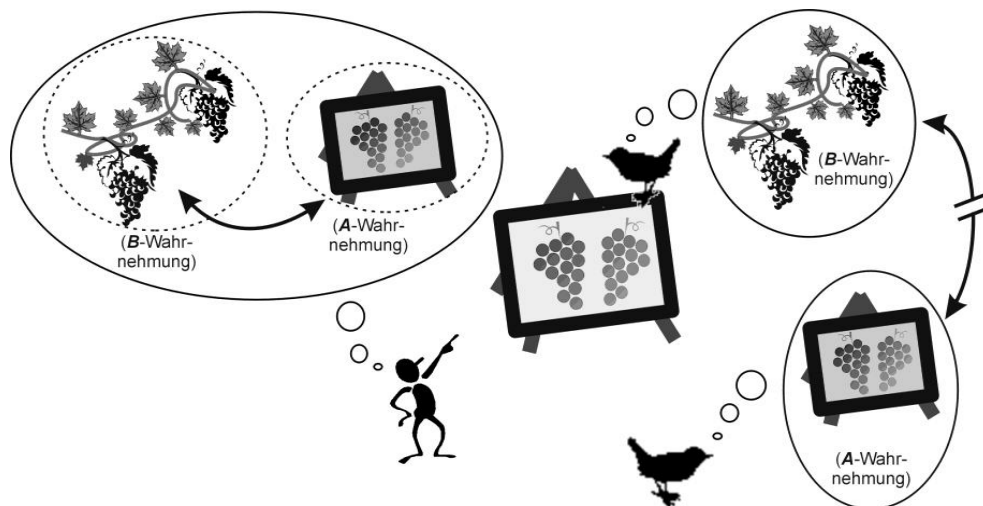


Abb. 1:
Erkennen von Ähnlichkeit: der dezeptive (rechts) und der immersive (links) Modus
(Die »Denkblasen« sollen nur jeweils die entsprechenden Verhaltensdispositionen graphisch-verkürzt andeuten).

Wegen dieser Einbettung kann allerdings der »Ähnlichkeitsträger« im immersiven Modus als Zeichenträger eingesetzt werden, um auf das aufmerksam zu machen, zu was er als ähnlich wahrgenommen wird. Einen Bildträger als Bild

²⁴ So definiert kann Ähnlichkeit, im Gegensatz etwa zu dem von N. Goodman angesetzten streng symmetrischen ontologischen Ähnlichkeitsbegriffs GOODMAN 1968, durchaus auch als unsymmetrische Relation aktualisiert sein.

²⁵ Bei Erschöpfung brechen die eingebetteten spontanen Reaktionen gelegentlich durch, oder wenn man sich sehr stark auf das Täuschungspotential einlässt, wie etwa beim Schauen spannender Spielfilme – das unwillkürliche, körperliche Erschrecken vor dem nur filmisch gezeigten Angriff eines großen Raubtiers – oder beim Computerspielen.

zu verwenden heißt also, eine durch den Bildträger auslösbare Täuschung als solche zu erkennen und den Bildträger – auch einem anderen gegenüber – als ein Zeichen zu benutzen, um auf die ›angetäuschte‹, sprich: ähnliche Situation zu verweisen.²⁶

Machen wir uns allerdings klar, dass ohne Kontextbildung der dezep-tive Modus zwar überwunden, d.h. der Zustand der Täuschung aufgehoben werden kann, dass dann aber nur der Zugang zu der *ohne* Verwechslung wahrgenommenen aktuellen Situation bleibt – ohne jede Beziehung zu der verwechselten Situation und mithin ohne Erkenntnis einer vorliegenden Ähnlichkeit (vgl. Abb. 1). So mag es etwa den Vögeln des Zeuxis gegangen sein (PLINIUS 1977), nachdem sie feststellten, dass das, was sie für essbare Früchte hielten, tatsächlich nur ein merkwürdiges flaches Brett ist. Demgegenüber bleibt die Illusion des Vorhangs vor Parrhasius' Bild für den getäuschten Zeuxis auch nach der Erkenntnis der Täuschung etwas, was ihn sich die Situation, in der ein Vorhang das von ihm eigentlich erwartete schlüpfrige Bild verdeckt, vorstellen lässt.

Diese handlungstheoretisch-semiotische Bildtheorie wird häufig im Gegensatz gesehen zu phänomenologischen Bildtheorien, in denen Bilder, statt als Zeichen, begriffen werden als Werkzeuge zum Herstellen einer besonderen Art von Gegenständen (WIESING 2005): Was durch ein Bild gegeben werde seien nämlich Gegenstände in ›reiner Sichtbarkeit‹ – ohne über das Visuelle hinausgehende Attribute.²⁷ Verwendungen als Zeichen seien dieser Primärfunktion nachzuordnen. Abgesehen davon, dass bei diesem Ansatz keinerlei intersubjektive Beziehung in Erscheinung tritt – die spezielle Wahrnehmungsform von Gegenständen in reiner Sichtbarkeit erfordert dergleichen nicht –, muss auch bezweifelt werden, ob die im Bildraum erscheinenden Gegenstände tatsächlich als Gegenstände ohne nicht-visuelle Eigenschaften erfasst werden. Das Problem besteht insbesondere darin, dass ein ›Haus‹ – um ein konkretes Beispiel zu nehmen – ohne Gewicht oder Undurchdringlichkeit seiner Wände (beides ganz klar nicht-visuelle Charakteristika unseres Alltagsbegriffs von Häusern) schlicht nicht mehr unter unsere übliche Unterscheidungsgewohnheiten für den Begriff ›Haus‹ fällt: Was immer in (oder eigentlich, als Werkzeug aufgefasst, ›mit‹) einem solchen Bild zu sehen wäre, es wäre gerade nicht mehr als Haus zu klassifizieren, bestenfalls noch als etwas, das ›visuell hausförmig‹ ist. Das entspricht keineswegs unserer Intuition zum Bildgebrauch: Ein gemaltes, gezeichnetes oder photographiertes Haus wird in aller Regel als ein Haus mit *allen* für Häuser typischen nicht-

²⁶ Auf die Besonderheiten der Ähnlichkeitskonstruktionen bei stark abstrahierten, logischen und reflexiv verwendeten (insbesondere ungegenständlichen) Bilder können wir hier aus Platzgründen nicht eingehen, verweisen aber auf die entsprechenden Ausführungen in SACHS-HOMBACH 2003 und SCHIRRA 2005.

²⁷ Begünstigt wurde diese Ablehnung zeichentheoretischer Bildtheorien sicher durch einen dort zunächst vorherrschenden semiotischen Zugang, der sich übermäßig an traditionellen Ansätzen zu propositionalen Zeichen orientierte und gleichzeitig ontologisch oder bewusstseinstheoretisch argumentierte: Das Spezifische der Bildverwendung gegenüber Sprache und ihre handlungstheoretische Einbettung waren dadurch weitgehend verdeckt.

visuellen Eigenschaften *in* einem Bild gesehen. Dieser Intuition kann die von uns vertretene modale Bildtheorie ganz ungezwungen folgen.

Im Kommunikationsakt mit wahrnehmungsnahen Zeichen wird eine Form der Indikation verwendet, die dem Zeigen bei sachbezüglichen Interaktionen recht nahe kommt: Genauer gesagt erlaubt es der in den immersiven Modus eingebettete dezeptive Modus, der ja mit einiger Wahrscheinlichkeit auch beim Rezipienten ausgelöst wird, die Ostentation des Kommunikators selbst konkret, d.h. körperlich nachzuvollziehen: Die mit der Verwechslung assoziierten Verhaltensdisposition werden bei *beiden* Kommunikationspartnern spontan ausgelöst. Beiden steht damit die tatsächlich nicht-anwesende Situation – wenn auch im Modus der Täuschung – »vor Augen«.

Als grundlegende Zeichenfunktion ergibt sich damit für den Gebrauch eines Bildes Folgendes: Ein Ensemble von Gegenständen, genauer: eine entsprechende Verhaltenssituation mit jenen Gegenständen, wird auf solche Weise interindividuell bereit gestellt, dass die Kommunikationspartner wenigstens für einige visuell überprüfbare Prädikationen die entsprechenden sensomotorischen Prüfverfahren in diesem Kontext unmittelbar vom Äußerungskontext aus anwenden (und kohärenztheoretisch auf Wahrheit überprüfen) können: Bildverwendung ist grundsätzlich empirische Kontextbildung.²⁸

Da es nicht das Bild ist, das etwas zeigt, sondern der »Sender« der sich mithilfe des Bildes als ein Wesen mit bestimmten Dispositionen darstellt (wiederum genauer: sich selbst gegenüber oder einem anderen), stellt sich die Frage nach einer präziseren Charakterisierung der Selbstdarstellung, die die Präsentation des Bildes notwendig begleitet. Unserer Auffassung nach sollte diese Ostentation der Bildkommunikationshandlung folgendermaßen formuliert werden: Ein Bild zu gebrauchen heißt, *»sich darstellen als einer, der dort etwas sieht, was gar nicht dort ist; der das auch weiß und der mit dieser Selbstdarstellung einem anderen gegenüber (oder sich selbst in der Rolle eines anderen) auf dieses zu Sehende aufmerksam macht, um auf diese Weise eine von der aktuellen abweichende Situation in den gemeinsamen Fokus der Aufmerksamkeit zu rücken«*.²⁹

Damit bilden bei der Kommunikation mit wahrnehmungsnahen Zeichen die ostentativen mit den indikatorischen Zeigeelementen eine charakteristische Verbindung: Während die sprachliche Bezugnahme auf eine andere als die Äußerungssituation den kontextuellen Bezugspunkt der Indikation schuldig bleiben muß, führt das Sich-Zeigen des Kommunizierenden im Falle von Bild- und anderen wahrnehmungsnahen Zeichen gerade zu einem vom Kommunikationspartner selbst wahrnehmbaren Zeigen der gemeinten Situa-

²⁸ Das schließt nicht aus, dass davon speziellere Verwendungsweisen abgeleitet werden können (vgl. SACHS-HOMBACH 2003).

²⁹ Es ist zu beachten, dass diese »Bedingung der Möglichkeit« von Bildern gleichwohl bei speziellen Bildgebräuchen (insbesondere reflexive Verwendungsweisen) in den Hintergrund treten oder sogar – im Sinne einer negativen Exemplifikation – außer Kraft gesetzt werden kann. Auf diese Weise lassen sich dann auch homogen-weiße Flächen als Bilder begreifen. Doch setzen derartige Verwendungsweisen stets den »darstellenden« Bildgebrauch als elementareren Bezugspunkt voraus.

tion. Der wesentliche Unterschied zwischen verbaler und pikturaler Kontextbildung besteht also darin, dass erstere die intendierte Situation nur *logisch* vergegenwärtigt, während letztere eine *empirische* Vergegenwärtigung ermöglicht: Mit Bildern wird ein Zugang zu einer anderen Situation eröffnet, bei dem die Wahrheit von – interindividuell laut geäußerten oder monologisch nur gedachten – Aussagen über die Situation in gewissen Grenzen unmittelbar empirisch überprüft werden kann.

3.2 Initiale Kontextbildung

Könnte ein Gedankenexperiment plausibel sein, in dem das Vermögen, situationsunabhängig mit Aussagen – zu kommunizieren, von den durch Bildern vermittelten Kontexten nicht nur ›unter anderem‹ Gebrauch macht, sondern überhaupt erst durch das täuschende Potential bildartiger Vorläufer möglich wird? Zwar bleibt die Verwendung eines wahrnehmungsnahen Zeichens mehrdeutig: Man kann ihm ja auch bloß im dezeptiven Modus gegenüber treten und nicht merken, dass es um einen *anderen* Kontext gehen soll. Aber dieser Mangel mag sich bei der begrifflichen Einführung der Kontextbildung als ein Vorteil entpuppen.³⁰

Zur methodischen Ableitung des Begriffs von Wesen, die zu Kontextbildungen fähig sind, stellen wir uns zunächst Wesen vor, die nur auf ihren je aktuellen Verhaltenskontext *a* Zugriff haben und daher einem (potentiellen) Ähnlichkeitsträger *B* höchstens im dezeptiven Modus – Verwechslung von Situation *a* mit einer Situation *b* – gegenüber treten können. Dieses Verhalten – etwa ein Beuteverhalten angesichts einer Attrappe – ist nicht notwendig auch kommunikativ. Zeichengebrauch läge erst vor, wenn das Wesen sich ändern gegenüber darstellen würde (Ostentation) als ein Wesen, das auf jene Stelle im aktuellen Kontext mit einem nicht adäquaten Verhalten, das für eine ganz andere Situation passend wäre (Indikation), reagiert oder zu reagieren bereit ist. Dazu muß es (i) anderen jenes Verhalten *vorführen*; (ii) auf dieses eigene Vorführverhalten auch *selbst* so reagieren, wie ein Empfänger, da seine Bedeutung ansonsten für Sender und Empfänger verschieden bliebe; und (iii) das Vorführen innerlich mit den dazu gehörigen Antwortreaktionen assoziieren (Triangulation).

³⁰ Ohne expliziten Bezug auf Kontextbildung haben I. Davidson und W. Nobel bereits 1989 eine ähnliche Überlegung vorgeschlagen (vgl. auch NOBLE/DAVIDSON 1996: Kap. 3).

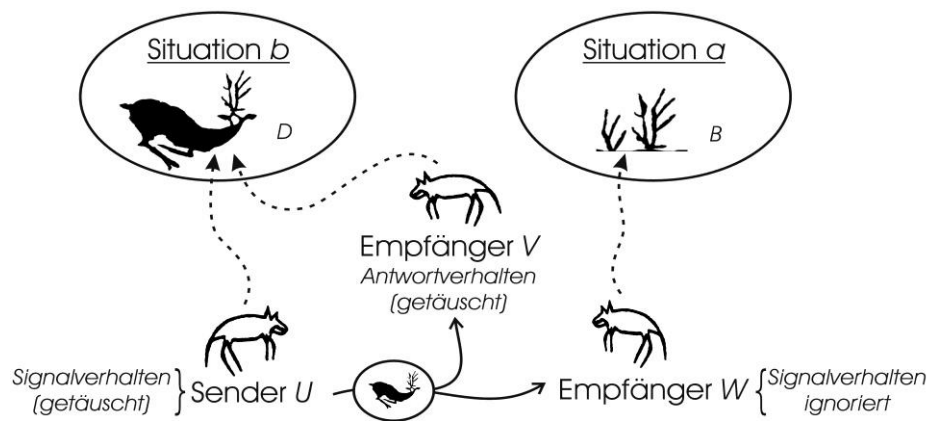


Abb. 2:
Schema zur Signalkommunikation in einer Verwechslungssituation (Hirsch/Geäst)

Die Besonderheit des kommunikativ eingesetzten dezeptiven Verhaltens ist nun, dass die Antwortreaktion zweigeteilt ausfallen kann, je nachdem, ob der Empfänger ebenfalls der Täuschung erliegt (Kontext *b*) oder nicht (Kontext *a*). Praktisch kann man sich durchaus vorstellen, dass etwa bestimmte natürliche Astformationen immer wieder zu Verwechslungen mit einer Beute Anlass geben. Signalsprachlich miteinander kommunizierende Gruppen könnten sich dort daran gewöhnen, entsprechende Signale eines Artgenossen in diesem Kontext zu ignorieren. Damit wäre ein Komplex aus einem Vorführverhalten (im dezeptiven Modus) und *zweierlei* Reaktionen darauf gegeben: Der Produzent reagiert mit seinem Signal auf den für ihn aktuellen Kontext *b*, auf den auch das Verhalten eines Rezipienten, der auf das Signal normal reagiert, gerichtet ist, während schließlich für einen Rezipienten, der das Signal auf die erlernte Weise ignoriert, Kontext *a* der aktuelle ist (Abb. 2). Auf der begrifflichen Ebene kann das Verinnerlichen des ursprünglich externen Signalverhaltens in einer solchen potentiellen Täuschungssituation entsprechend *beide* Reaktionsoptionen bei dem reflexiven Produzenten intern auslösen und so das Fundament für eine Relation zwischen den beiden beteiligten Kontexten für Wesen dieser Art bilden. Es findet eine doppelte semiotische Triangulation statt: Einerseits zwischen Produzent, Rezipient 1 und Zeichenträger hinsichtlich der Ostentation (›wir reagieren auf den Zeichenträger auf bestimmte Weise« – nämlich getäuscht-inadäquat), und andererseits zwischen Produzent, Rezipient 2 und ›angetäuschter« Situation hinsichtlich der Indikation (›wir reagieren auf die dort anwesenden Gegenstände auf bestimmte Weise«).

In dem Gedankenexperiment können wir daher davon sprechen, dass sich Wesen, die unter einen solchen Begriff fallen, einander (und sich selbst gegenüber) darstellen als Wesen, die etwas wahrnehmen, was so gar nicht anwesend ist; die weiterhin annehmen, dass auch die jeweiligen Kommunikationspartner sich gleichermaßen so darstellen können. Dabei bleibt dieses Verhalten allerdings noch ganz abhängig von der wirklichen Anwesenheit des Gegenstands *B* im aktuellen Kontext, einem Gegenstand, der also tatsächlich gerade wahrgenommen werden kann und dabei wegen entsprechender spon-

tan ausgelöster Reaktionen leicht mit etwas anderem (*D*) zu verwechseln (d.h. ihm ähnlich) ist.

4. Die Fähigkeit, sich als Bildbenutzer darstellen zu können

Das Gedankenexperiment zur initialen Kontextbildung führt zum Begriff eines Wesens, dem eine Vorform der piktorialen Kontextbildung gelingt: Es hat den Bildträger im immersiven Modus gebraucht und dabei eine Ostentation mit einer Indikation auf eine nicht anwesende Situation verbunden. Allerdings lässt sich von außen her (d.h. insbesondere aus der Perspektive des empfangenden Kommunikationsteilnehmers) nur vermuten, dass mehr als eine einfache Verwechslung – der reine dezeptive Modus also – die Grundlage eines damit inadäquat abgegebenen Signals gewesen ist. Ebenso bleibt für den sendenden Part ohne weiteres unklar, ob sein Gegenüber die Kontextbildung nachvollzogen hat, oder auf eine Ähnlichkeit lediglich im Modus des Getäuscht-Seins reagiert. Die doppelte Triangulation schafft zwar den logischen Raum für die piktoriale Kontextbildung; doch sind weitere kommunikative Absicherungen notwendig, um die damit gegebene Überschreitung des Hier und Jetzt als gemeinsam und damit wechselseitig kontrolliert und wiederholbar zu stabilisieren.

4.1 Kommunikative Stabilisierung des Bildgebrauchs

Sofern wir annehmen könnten, dass die beteiligten Kommunizierenden sich zusätzlich darüber verständigen, dass es sich bei ihrer aktuellen Interaktion nicht um eine einfache Verwechslung handelt, sondern um eine *absichtsvolle* Nutzung/Artikulation von Ähnlichkeit, wäre zumindest ein gewisser Ansatz gegeben, den immersiven vom rein dezeptiven Modus beim jeweiligen Gegenüber zu unterscheiden. Für eine solche herausgehobene Verwendungsweise des Zeichens gegenüber dem Ähnlichkeitsträger könnte etwa das bei höheren Säugetieren verbreitete ›Spielgesicht‹ als zusätzliche ›Immersions-Markierung in die Gesamtkommunikation eingehen.³¹

Beachten wir aber, dass in der Situation der initialen Kontextbildung offensichtlich bereits *zwei* Zeichen in Kombination verwendet worden sind: Der Ähnlichkeitsträger nämlich *und* das Signal, das sich auf die eigentlich nur ähnliche Situation bezog und in der tatsächlich aktuellen Situation unpassend war – eine (noch ungesicherte) piktoriale Kontextbildung *und* eine Zeichen-

³¹ Analog wäre auch das ›Entwarnungssignal‹ (s.o.) als Markierung dieser Komponente des Selbstbezugs denkbar. Unter ›Spielgesicht‹ wird eine bei Säugetieren verbreitete Signalart zusammengefasst, mit der verhindert wird, dass spielerisch-aggressives Verhalten von Jungtieren (z.B. gegenüber Erwachsenen) zu ernsthaften Aggressionen ausartet. Das »Spielgesicht«, das mit dem auch für menschliche Kommunikationsbereitschaft assoziierten »Öffnen des Gesichts« in Beziehung steht, signalisiert nach Eibl-Eibelsfeld (1997: 190) eine »freundliche Beißintention«.

verwendung, die sich auf die so evozierte Situation bezieht. Es ist, wie sich zeigt, diese Verschränkung *beider* Zeichenhandlungen, die letztlich eine Stabilisierung des immersiven Modus zur Bildzeichenhandlung einerseits und zur Ablösung sprachlicher Zeichen vom strikten Bezug auf die Äußerungssituation andererseits ermöglicht.

Da wir für das Begriffsfeld der Wesen mit initialer Kontextbildung bereits die semiotische Triangulation in Anspruch genommen haben, können wir ohne Beschränkung der Allgemeinheit davon ausgehen, dass auch die syntaktisch gegliederten, selbstreferentiellen, intersubjektiv kontrollierten ästhetischen Produktions-/Rezeptions-Situationen, wie M. Vogel (2001: § 4.2.2.3) sie herausgearbeitet hat, bei den betrachteten Wesen gegeben sind. Im Sinne eines Spiels sind Wesen, die unter diesen Begriff fallen, also in der Lage, sich gegenseitig Variationen komplexer Verhaltensweisen vorzuführen, deren Sinn noch nicht jenseits der sozialen Handlung selbst und den darin gegebenen Variationsmöglichkeiten liegt.³²

Des weiteren können wir annehmen, dass auch das kommunikative Handeln mit Sachbezug bei den Wesen im Gedankenexperiment über das Stadium einfacher Signale hinausgehen kann und subordinierend koordinierte Verweise auf mehrere Eigenschaftsdimensionen an gegenwärtigen (Prä-)Objekten artikuliert werden können, wie Arno Ros (1979: 77f.) erläutert hat. Es können also – nach wie vor strikt situationsgebundene – Vorformen der Differenzierung einer Äußerung in nominatorische und prädikatorische Anteile auftreten, die sich gleichwohl nur auf aktuell (im Fall einer Täuschung auch nur scheinbar) Wahrgenommenes beziehen³³ (analog etwa zu: »(etwas) Rundes ist (hier), das (zugleich) glänzend ist« und »das Runde (hier) (das gerade noch glänzend war) glänzt (jetzt) nicht (mehr)«).³⁴ Dabei liegt es nahe, dass die gesteigerte Fähigkeit zur spielerischen Nutzung syntaktisch komplexer Interaktionen auch zum Ausdruck der gewachsenen Differenzierungen der Kommunikationsmöglichkeiten angewendet wird.³⁵

Übertragen auf die Situation der initialen Kontextbildung können wir also durchaus davon ausgehen, dass nicht nur ein monolithisches Signal im Sinne der oben erwähnten einfach klassifikatorischen Warnrufe relativ zu dem Ähnlichkeitsträger verwendet wird, sondern ein syntaktisch bereits gegliedertes Zeichen, das, obschon an die aktuelle Situation gekettet, zudem

³² Im Übrigen ist auf diese Weise gegeben, dass nicht notwendig auf eine vorgefundene Umweltformation als Ähnlichkeitsträger gewartet werden muß: Eine im Verlauf einer solchen Variationsreihe erzeugte Formation mag ebenfalls als Auslöser eines dezeptiven Modus geeignet sein. Sie steht damit am Anfang der Genese der Fähigkeit, selbst Bildträger mit Absicht zu erzeugen.

³³ Wegen der unauflösbaren Bindung an die jeweilige Äußerungssituation bleiben derartig syntaktisch gegliederten Zeichen mit subordinierten Inhaltsspielräumen Quasiprädikationen im Sinne Tugendhats (1976: §§ 13 & 19).

³⁴ Die in Klammern gesetzten Äußerungsteile sind nur der besseren Verständlichkeit wegen angegeben – auf der angenommenen Kommunikationsstufe sind ihre Äquivalente tatsächlich noch nicht möglich.

³⁵ Mithin ist denkbar, dass eine zunächst unabhängig von dem Teile der Umgebung darstellenden Kommunikationsverhalten gebildete, nämlich im sozialen Spiel etablierte Syntax nachträglich in den Dienst der kommunikativen Darstellungsfunktion genommen wird.

den Vorläufer einer funktionalen Differenzierung in nominatorische und prädikatorische Aspekte aufweist. Wesentlich ist hierbei, dass damit in der Situation der initialen Kontextbildung *mehrere* Vorkommen eines unter denselben Typus fallenden Gegenstands (»etwas Rundes«) mit verschiedenen Ausprägungen in einer subordinierten Eigenschaftsdimension (»glänzend«/»nicht-glänzend«) als in zweierlei Situationen einander kommunikativ gegenübergestellt werden können. Die gemeinsame *Identifizierung* als Erscheinungsweisen *eines* Gegenstands ist dabei die Kehrseite der gemeinsamen *Differenzierung* in *verschiedene* Kontexte.

Damit ergibt sich der Bildgebrauch als abhängig von einer komplexen und potentiell offenen Folge von abwechselnden Zeichenverwendungen mit jeweils tatsächlich *drei* Komponenten: (i) der Ostentation gegenüber dem Ähnlichkeitsträger als piktoriale Kontextbildung, (ii) der syntaktisch bereits differenzierten Quasiprädikation (als Indikation) und (iii) der Immersionsmarkierung durch das Spielgesicht im Wechsel mit einer variierenden syntaktisch differenzierten Quasiprädikation relativ zu einem tatsächlich anwesenden entsprechenden Gegenstand, die ohne Immersionsmarkierung geäußert wird. Das Äußerungsspiel, dass sich damit spielen läßt, indem zwei Kommunikationspartner einander mal die eine, mal die andere Erscheinungsform – einmal mit Immersionsmarkierung und der einen subordinierten Eigenschaft, einmal ohne Spielgesicht und mit der anderen subordinierten Eigenschaft – signalisieren, bildet die dialogische Grundlage, auf der Kontextbildung und Objektkonstitution kommunikativ stabilisiert und damit echter Bildgebrauch *und* voll-propositionale Sprache möglich werden.

4.2 Kontextbildung ohne wahrnehmungsnahes Medium

Bleibt die Frage, wie die nach wie vor gegebene Fixierung der Fähigkeit zur Kontextbildung auf einen *anwesenden* Ähnlichkeitsträger – nun zum echten Bildträger geworden – aufgehoben werden kann, so dass auch rein logische Kontextbildungen begrifflich rekonstruiert wären: Wie kann, mit anderen Worten, eine Ostentation relativ zu der lokal gar nicht adäquat durchführbaren Indikation auf eine nicht einmal täuschend-stellvertretend präsente Situation – ohne Beteiligung eines wahrnehmungsnahen Zeichens – erreicht werden? Es ist ja der faktisch gegenwärtige Ähnlichkeitsträger, der es überhaupt erst möglich gemacht hat, die Kontextbildung als intersubjektives Unternehmen begrifflich einzuführen.

Könnte diese Fähigkeit wiederum durch eine Art ostentative Selbsttäuschung erklärt werden? Die Selbstdarstellung als ein Wesen, das sich relativ zur *tatsächlichen* Situation *inadäquat* darstellt, ist bereits in der Situation der initialen Kontextbildung ein essentieller Bestandteil. Im Fall der rein logischen Kontextbildung wäre eine doppelte solche Selbstdarstellung denkbar: Sich darstellen als ein Wesen, das sich über die Anwesenheit eines wahr-

nehmungsnahen Zeichens in der tatsächlichen Kommunikationssituation täuscht – mit allen Auswirkungen, die das auf seinen Leib und den damit gegebenen Wahrnehmungs- und Ausdrucksvermögen bis hin zur empirischen Kontextbildung in einer Kommunikationsdyade hat; das sich, kurz gesagt, inadäquater Weise so verhält, als wäre es in einer Situation der initialen Kontextbildung (vgl. Abb. 3).

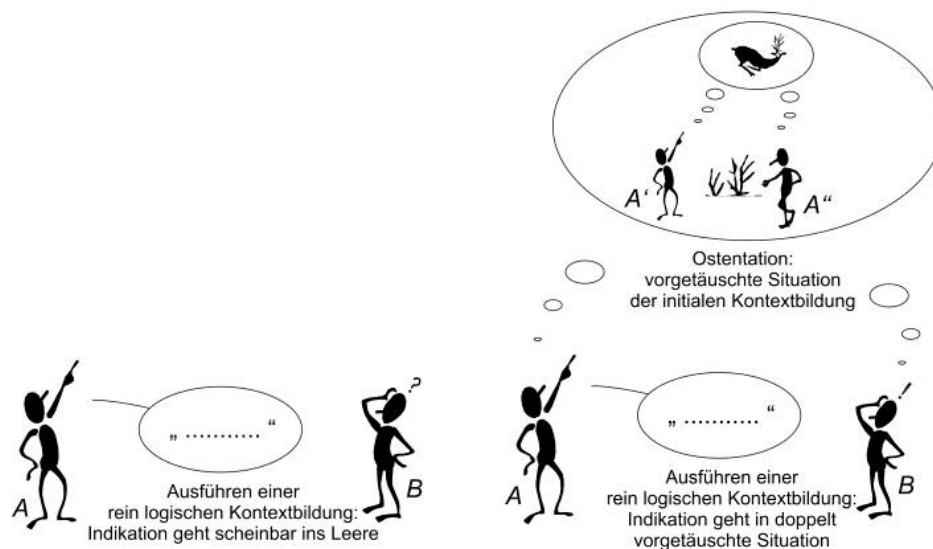


Abb. 3:
Indikation und Ostentation bei rein logischer Kontextbildung
Die »Denkblasen« sind wieder als graphische Verkürzungen handlungstheoretischer Dispositionen zu verstehen.

Die Vermittlung in diesem zweiten Gedankenexperiment – der »Ähnlichkeitsträger«, an dem nun die Täuschung festgemacht wird – ist nichts anderes als das Wesen selbst, das sich auch für sich selbst »so ähnlich verhält« wie ein Wesen, das sich tatsächlich in einer Situation empirischer Kontextbildung befindet. So, wie zuvor ein – ursprünglich in einem Signal benutztes – Verhalten, ausgelöst von einem tatsächlich anwesenden Gegenstand, auf eine *andere*, äußere Realität ausgerichtet wird, so richtete sich nun, ausgelöst von der tatsächlich wirksamen Selbstvergegenwärtigung körperlicher und psychischer Zustände, das Zeichenverhalten auf eine *andere*, »innere« Realität: die nämlich als eine körperliche Selbstvergegenwärtigung der eigenen Zustände für eine andere Verwechslungssituation – der Situation der zugehörigen initialen Kontextbildung mit einem passenden Ähnlichkeitsträger – zu charakterisieren wäre. Das Gedankenexperiment liefert also tatsächlich eine mögliche Erklärung einer von der Anwesenheit eines wahrnehmungsnahen Zeichens unabhängigen, nicht-empirischen, d.h. einer nun rein logischen Form der Kontextbildung. Das benötigte wahrnehmungsnahе Zeichen wandert gewissermaßen mit einer verinnerlichten Ostentation *in* das kontextbildende We-

sen selbst. Dieses wird sich – verkürzt gesprochen – selbst zum ›Bild‹.³⁶ Das kommunikative Gegenüber muß wiederum aus dem situativ zwar inadäquaten, aber mit einer Situation der initialen Kontextbildung kompatiblen und in sich schlüssigen komplexen Kommunikationsverhalten erschließen, dass der so Kommunizierende ihm gegenüber gerade eine Kontextbildung ohne stützendes wahrnehmungsnahes Zeichen vollziehen will, wobei es im Wesentlichen selbst eine entsprechende Selbstdarstellung mit scheinbarer (›innerer‹) piktorialer Kontextbildung nachvollzieht.

4.3 Die Fähigkeit, sich etwas visuell vorzustellen

Interessanter Weise kommen Menschen, die an einer rein sprachlichen Kontextbildung beteiligt sind, in *Erklärungen* ihrer Kommunikationshandlungen häufig auf ihr Vorstellungsvermögen zu sprechen. Sie reden sogar davon, ›innere Bilder‹ zu benutzen, welche als Fundament der sprachlichen Kontextbildung dienen. Der sprachliche Verweis auf ein rein hypothetisch angesetztes Bild – »im mentalen Bild, das ich gerade vor meinem inneren Auge habe, (... ist dies und das der Fall ...)« – scheint ähnlich zu funktionieren wie die sekundären verbalen Kontextbildungen durch Verweis auf ein (abwesendes) wahrnehmungsnahes Zeichen – »in Caravaggios Bild ›Bacchus‹ (... ist dies und das der Fall...)«.

Nur kann es sich dabei nicht um etwas handeln, das tatsächlich unter den Begriff ›Bild‹ in dem Sinne, wie er in Abschnitt 3 näher bestimmt wurde, fällt. Jedenfalls läßt sich kein materieller (oder apparativer) Bildträger zuordnen, der mehreren Bildnutzern in einer gemeinsamen Zeichenhandlung zugänglich wäre. Der zentrale Punkt der interindividuellen Verfügbarkeit wird hier ausdrücklich ausgeschlossen. Aus diesem Grund bevorzugen wir die in dieser Hinsicht weniger problematische Rede vom ›visuellen Vorstellungsvermögen‹ gegenüber der vom ›Haben mentaler Bilder‹.³⁷

Dass sich jemand etwas visuell vorstellt, erfahren wir dadurch, dass jene Person uns gegenüber *davon spricht*. Sie stellt sich dar – und zwar auch sich selbst gegenüber – als jemand, der etwas visuell wahrnimmt und sich entsprechend verhält, obwohl es nicht gegenwärtig ist. Sie stellt sich etwa dar als eine Person, die die Front ihres Wohnhauses sieht und dabei die Fenster zählt, obwohl diese Front gerade gar nicht gesehen werden kann – ähnlich also der Situation, in der sie ein entsprechendes Bild des Hauses verwendet; nur gibt es im aktuellen Verhaltenskontext keinen zugehörigen Bildträger. Oder es handelt sich um jemanden, der einer Fußball-Live-Reportage im Radio zuhört und sich die sportlichen Ereignisse ›vor Augen führt‹: der sich also

³⁶ Beltings Bildanthropologie (2001) hat dieses Moment, wenn auch nicht in begriffsgenetischem Zusammenhang, verdeutlicht.

³⁷ Es wäre ein allzu offensichtlicher Reifizierungsfehler, wenn man von einem Handlungsaspekte erläuternden Begriff, der statt durch »sich etwas visuell vorstellen« manchmal auch durch die Phrase »ein mentales Bild von etwas haben« ausgedrückt wird, einfach zur empirischen Untersuchung einer Gegenstandssorte »mentale Bilder« überginge; eine der vielen Verhexungen unseres Verstandes durch bestimmte Aspekte unserer Sprache.

– für sich selbst – darstellt als jemand, der jenem Geschehen wahrnehmend beiwohnt und nicht bloß einen situativ rein logisch verankerten sprachlichen Bericht davon bekommt (vgl. SCHIRRA 1995). Man mag also, zumindest angesichts der beiden Beispielsituationen, durchaus annehmen, dass es die Situation einer piktorial-empirischen Kontextbildung ist, in der sich zu befinden jene Person durch ihr sprachliches Verhalten vortäuscht, wenn sie von ihrem Vorstellungsvermögen redet.

Die Fähigkeit, sich etwas anschaulich vorzustellen, tritt also als ein Begriff auf, mit dessen Hilfe Aspekte einer reflexiven Kommunikation unterschieden werden: Jemand stellt sich *sich selbst gegenüber* dar als ein Wesen mit bestimmten Handlungsperspektiven. Dabei wird auf diese Selbstdarstellung verwiesen, um verständlich zu machen, weshalb eine Kontextbildung ohne anwesenden Ähnlichkeitsträger erster Stufe bewerkstelligt werden kann. Es geht also weniger darum, dass die scheinbare Bildwahrnehmung ein primäres Ereignis ist, dem – kontingenter Weise – eine Erläuterung folgt. Vielmehr tritt die vorgetäuschte Darstellung einer Bildwahrnehmung nur zu dem Zweck der Erläuterung des Funktionierens rein logischer Kontextbildungen überhaupt auf – sowohl anderen wie sich selbst gegenüber. Auch hierbei ist die reflexive Selbstvergegenwärtigung eigentlich abgeleitet von einer Selbstdarstellung einem anderen gegenüber.

Wer sich etwas anschaulich vorstellt, stellt sich mithin nicht nur als jemand dar, der eine entsprechende piktorale Kontextbildung vortäuscht: Seine Ostentation ist vielmehr auf die Situation der *initialen* Kontextbildung gerichtet und damit auf eine Erläuterung im Sinne einer Begriffsgenese, wie es zur Fähigkeit der Kontextbildung gekommen ist. Ist dieses Erklärungsschema erst einmal etabliert, kann es als ein Bezugspunkt für logische Kontextbildungen auch für den Fall dienen, in dem tatsächlich gar kein visuelles Vorstellungsvermögen – keine komplexe reflexive Ostentation mit vorge-täuschter initialer Kontextbildung – verwendet wird.

Diese »referentielle Grundierung« der logischen Kontextbildung löst dabei in gewisser Weise auch das in Abschnitt 2.3 erwähnte Referenzproblem: Werden in der bewusstseinstheoretisch orientierten Kognitionswissenschaft mentale Bilder noch als spezielle geistige Entitäten betrachtet, die an Stelle von nicht-vorhandenen Perzeptionen eine referentielle Verankerung sprachlicher Ausdrücke »in der Welt« übernehmen sollen, falls über in der Äußerungssituation Abwesendes (etwa *x*) geredet wird, so verschiebt sich der Fokus im handlungstheoretisch-begriffsgenetisch verstandenen Begriff des »Habens mentaler Bilder von *x*« bzw. besser: der »Fähigkeit, sich *x* visuell vorstellen zu können« darauf, die Fähigkeit zur rein logischen Kontextbildung zu rekapitulieren. Es geht dabei also gar nicht um eine unmittelbare Beziehung zwischen Sprache und außersprachlicher Bedeutung, sondern vielmehr um die Form der Erläuterung einer solchen Beziehung als Aspekte komplexer Kommunikationshandlungen.

Bei genauerem Hinsehen wird der Zusammenhang zwischen Vorstellungsvermögen und Kontextbildung im Übrigen auch grammatisch deutlich.

Dabei scheinen die beiden Arten von »sich x vorstellen« zunächst auf andere Zeichenfunktionen hinzuweisen: Während die propositionale Form »sich visuell vorstellen, dass p « – etwa »sich visuell vorstellen, dass die Königin von Großbritannien kahl sei« – auf die verinnerlichte Selbstdarstellung analog zum Gebrauch einer auf Visuelles bezogenen Aussage zu funktionieren scheint, evoziert die substantivische Form »sich n visuell vorstellen« – etwa »sich den Präsidenten der französischen Republik visuell vorstellen« die Selbstdarstellung als jemand, der eine verinnerlichte Nomination vollzieht. Allerdings sollte klar sein, dass sowohl Nomination wie Proposition auch bei verinnerlichter Verwendung ohne Kontextbezug nicht funktionieren können.³⁸ Ein Gegenstand tritt uns konkret nie isoliert, sondern stets als Teil bestimmter Kontexte in Erscheinung, so dass es bei der substantivischen Form nur darum gehen kann, sich einen Kontext mit jenem Gegenstand vorzustellen. Auch eine Proposition kann nicht abgetrennt von einer Situation betrachtet werden, in der etwa die jeweiligen Implikaturen und Präsuppositionen ebenfalls gelten. Auch hier markiert das Zuschreiben einer visuellen Vorstellung also letztlich das Etablieren eines Kontextes, in dem jene Proposition gilt.

5. Resümee

Die anthropologische Funktion von wahrnehmungsnahen Zeichen – allen voran von Bildern – liegt, so haben die beiden Gedankenexperimente nahe gelegt, in der Vermittlung des Begriffs der Fähigkeit zur Kontextbildung. Dabei geht es nicht um eine empirische Bestimmung des prähistorischen Geschehens, sondern um das auf unsere Unterscheidungsgewohnheiten gerichtete Verstehen dessen, was alles mit dem Begriff einer Kompetenz zu kommunikativem Verhalten, mit dem der gemeinsame Interessenfokus auf eine abwesende Situation ausgerichtet werden kann, vorausgesetzt werden muß und welche Begriffe den einzuführenden Begriff selbst bereits voraussetzen.

Diese Funktion der Bilder beruht wesentlich darauf, dass die beiden beim Kommunizieren notwendig involvierten Arten des Zeigens bei Bildkommunikation in einem besonderen Verhältnis zueinander stehen: Wie bei den einfacheren sachbezüglichen Interaktionen orientiert sich nämlich das Verhalten der beiden Interaktionspartner gleichermaßen an aktuell Wahrnehmbarem. Allerdings handelt es sich dabei um eine Wahrnehmung im dezeptiven Modus – im Zustand der Täuschung. In der Bildverwendung ist diese – in der Situation eigentlich inadäquate – Indikation aber durch eine ostentative Selbstdarstellung des Bildverwenders als ein sich bewusst Täuschender

³⁸ Abgesehen davon zeigen bereits leichte Variationen der Beispielsätze, dass eine strikte Interpretation als Nomination oder Proposition ohnehin nicht tragfähig ist: »sich einen Apfel visuell vorstellen« und »sich visuell vorstellen, dass der König von Frankreich kahl sei«. Im ersten Fall handelt es sich nicht um eine Nomination, sondern um eine verkappte Existenzaussage, im zweiten um eine Proposition mit unspezifisch-fiktivem Gehalt. Beide sind wesentlich besser als Verweise auf vorgetäuschte empirische Kontextbildungen zu begreifen.

kommunikativ nutzbar gemacht: Statt einer unbewussten Täuschung zu erliegen wird nun im immersiven Modus ein Fall von Ähnlichkeit erkannt.

Begriffsgenetisch wird der Boden für eine potentiell intersubjektive Kontextbildung im immersiven Modus – d.h. eine initiale Kontextbildung mit Hilfe eines Ähnlichkeitsträgers – durch eine doppelte semiotische Triangulation bereitet, wie sie nur relativ zu den zwei möglichen Reaktionen auf ein in dezeptivem Modus geäußertes Signal erfolgen kann: Im Bildgebrauch müssen die Verhaltensweisen sowohl auf den Bildträger wie auf den Bildinhalt miteinander in Beziehung stehen. Doch erst die Einbettung solcher ersten pikturalen Kontextbildungen in weitere komplexe quasi-propositionale Zeichenhandlungen stabilisiert den gemeinsamen Zugang der beiden Kommunikationspartner zu einer abwesenden Situation und eröffnet zugleich den Weg zu einer echt propositionalen Kommunikation sowie zum Umgang mit individuierten persistenten (sortalen) Gegenständen.

Die anthropologische Rolle der Bilder ist allerdings nicht auf die initiale Kontextbildung – die begriffliche Einführung der Kontextbildung in Gegenwart eines Ähnlichkeitsträgers – beschränkt. Auch die Fähigkeit zu rein logischen Kontextbildungen hängt letztlich am Bildbegriff: Denn es ist eine selbst-täuschende Vergegenwärtigung der Situation der initialen Kontextbildung, auf die zur Erklärung einer nicht durch einen täuschenden Gegenstand ausgelösten Kontextbildung verweist, wer einer Person (auch sich selbst) den Gebrauch des (visuellen) Vorstellungsvermögens im Zusammenhang mit der Verwendung einer rein logischen (sprachlichen) Kontextbildung zuschreibt. Auf diese Weise macht man in einer Situation, in der eigentlich keine andere Beziehung zu der gemeinten nicht-anwesenden Situation besteht, auf die begriffsgenetische Rekonstruktion des Begriffsfeldes von Wesen, die die Fähigkeit zur Kontextbildung durch Bildgebrauch besitzen, aufmerksam: Obwohl keine Bilder beteiligt sind, reden wir dann unser Verhalten erläuternd häufig von ›inneren Bildern‹, die wir hätten.

Gleichsam als ›Geburt der Kontextbildung aus dem Geist des Bildes‹ stellt die philosophische Betrachtung somit den Bildbegriff an den entscheidenden anthropologischen Wendepunkt neben den Sprachbegriff: Denn ohne Bildkompetenz wäre keine Befreiung aus den Bedingungen des Hier und Jetzt der je aktuellen Verhaltenssituationen zu erreichen. Der Ursprung des Bildes ist der Ursprung des Menschen.

Literatur

BELTING, HANS: *Bild-Anthropologie. Entwürfe für eine Bildwissenschaft*. München [Fink] 2001

DAVIDSON, DONALD: *Subjective, Intersubjective, Objective*. Oxford [Clarendon Press] 2001

- DAVIDSON, IAIN; WILLIAM NOBLE: The Archaeology of Perception. Traces of Depiction and Language. In: *Current Anthropology*, 30, 1989, S. 125-155
- EIBL-EIBLSFELD, IRENÄUS: *Die Biologie des menschlichen Verhaltens. Grundriß der Humanethologie*. 3. Auflage. Weyarn [Seehamer] 1997
- GOODMAN, NELSON: *Languages of Art. An Approach to a Theory of Symbols*. Indianapolis [Bobbs-Merill] 1968
- LAKOFF, GEORGE: The Invariance Hypothesis: Is Abstract Reason Based on Image-Schemas? In: *Cognitive Linguistics*, 1, 1990, S. 39-74
- LORENZ, KUNO: *Einführung in die philosophische Anthropologie*. Darmstadt [WBD] 1990
- MEAD, GEORGE HERBERT: *Mind, Self, and Society*. Chicago [U of Chicago P] 1934
- NOBLE, WILLIAM; IAIN DAVIDSON: *Human Evolution, Language, and Mind. A Psychological and Archaeological Inquiry*. Cambridge [Cambridge UP] 1996
- PLESSNER, HELMUTH: *Die Stufen des Organischen und der Mensch*. Berlin [Götschen] 1928
- PLINIUS SECUNDUS D.Ä.: *Naturkunde (Naturalis Historiae)*. Lat.-dt. Buch XXXV: *Farben, Malerei, Plastik*. Herausgegeben und übersetzt von Roderich König in Zusammenarbeit mit Gerhard Winkler. 2., überarbeitete Auflage. Düsseldorf [Artemis & Winkler] 1977
- ROS, ARNO: *Objektkonstitution und elementare Sprachhandlungsbegriffe*. Königstein Ts. [Hain] 1979
- ROS, ARNO: Was ist Philosophie? In: RAATZSCH, RICHARD (Hrsg.): *Philosophieren über Philosophie*. Leipzig [Leipziger Universitätsverlag] 1999, S. 36–58
- ROS, ARNO: *Materie und Geist. Eine philosophische Untersuchung*. Paderborn [Mentis] 2005
- SACHS-HOMBACH, KLAUS: *Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft*. Köln [Halem] 2003
- SCHIRRA, JÖRG R.J.: Understanding Radio Broadcasts On Soccer: The Concept »Mental Image« and its Use in Spatial Reasoning. In: SACHS-HOMBACH, KLAUS (Hrsg.): *Bilder im Geiste: Zur kognitiven und erkenntnistheoretischen Funktion piktoraler Repräsentationen*. Amsterdam [Rodopi] 1995, S. 107-136
- SCHIRRA, JÖRG R.J.: *Foundation of Computational Visualistics*. Wiesbaden [DUV] 2005
- SCHIRRA, JÖRG R.J.; KLAUS SACHS-HOMBACH: Fähigkeiten zum Bild- und Sprachgebrauch. In: *Deutsche Zeitschrift für Philosophie*, 54, 2006, S. 887-905
- SCHIRRA, JÖRG R.J.; MARTIN SCHOLZ: Zwei Skizzen zum Begriff »Photorealismus« in der Computergraphik. In: SACHS-HOMBACH, KLAUS; KLAUS REHKÄMPER (Hrsg.): *Bild – Bildwahrnehmung – Bildverarbeitung. Interdisziplinäre Beiträge zur Bildwissenschaft*. Wiesbaden [DUV] 1998, S. 69-79

- TUGENDHAT, ERNST: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*. Frankfurt/M. [Suhrkamp] 1976
- VARELA, FRANCISCO J.; EVAN THOMPSON; ELEANOR ROSCH: *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, MA [MIT Press] 1991
- VOGEL, MATTHIAS: *Medien der Vernunft. Eine Theorie des Geistes und der Rationalität auf Grundlage einer Theorie der Medien*. Frankfurt/M. [Suhrkamp] 2001
- WHEELER, MICHAEL: *Reconstructing the Cognitive World. The Next Step*. Cambridge, MA [MIT Press] 2005
- WIESING, LAMBERT: Medienphilosophie des Bildes. In: SANDBOTHE, MICHAEL; LUDWIG NAGL (Hrsg.): *Systematische Medienphilosophie*. Berlin [Akademie] 2005, S. 147-162
- WITTGENSTEIN, LUDWIG: *Tractatus logico-philosophicus*. London [Kegan Paul] 1922