

Gruppen zur Herstellung von Zugehörigkeit bzw. zur Abgrenzung nach außen eingesetzt werden.

### Auswahlbibliographie

- Inci DIRIM, Peter AUER: Türkisch sprechen nicht nur die Türken – Über die Unschärfebeziehung zwischen Sprache und Ethnie in Deutschland. Berlin 2004.
- Aldo DI LUZIO, Susanne GÜNTNER, France ORLETTI (Hgg.): Culture in Communication. Amsterdam 2001.
- Alessandro DURANTI: Linguistic Anthropology. Cambridge 1997.
- Alessandro DURANTI: Linguistic Anthropology: History, Ideas, and Issues. In: Alessandro: DURANTI, Linguistic Anthropology. A Reader. Malden MA 2001, S. 1-38.
- William A. FOLEY: Anthropological Linguistics. An Introduction. Malden MA / Oxford 1997.
- Susanne GÜNTNER: Diskursstrategien in der Interkulturellen Kommunikation. Analysen deutsch-chinesischer Gespräche. Tübingen 1993.
- Susanne GÜNTNER: Zur Aktualisierung kultureller Differenzen in Alltagsinteraktionen. In: Stefan RIEGER, Schamma SCHAHADAT, Manfred WEINBERG (Hgg.): Interkulturalität. Zwischen Inszenierung und Archiv. Tübingen 1999, S. 251-268.
- Susanne GÜNTNER: Eine Sprachwissenschaft der ‚lebendigen Rede‘. Ansätze einer anthropologischen Linguistik. In: Angelika LINKE, Hanspeter ORTNER, Paul R. PORTMANN-TSELIKAS (Hgg.): Sprache und mehr. Ansichten einer Linguistik der sprachlichen Praxis. Tübingen 2003, S. 189-209.
- Susanne GÜNTNER, Thomas Luckmann, Thomas: Wissensasymmetrien in der interkulturellen Kommunikation. In: Helga KOTTHOFF (Hg.): Kultur(en) im Gespräch. Tübingen 2002, S. 213-242.
- William F. HANKS: Discourse Genres in a Theory of Practice. In: American Ethnologist 14/4 (1987), S. 668-692.
- Inken KEIM: Kommunikative Praktiken in türkischstämmigen Kinder- und Jugendgruppen in Mannheim. In: Deutsche Sprache 32/3 (2004), S. 198-226.
- Thomas LUCKMANN: Kommunikative Gattungen im kommunikativen ‚Haushalt‘ einer Gesellschaft. In: Gisela SMOLKA-KOERDT, Peter SPANGENBERG, Dagmar TILLMANN-BARTYLLA (Hgg.): Der Ursprung der Literatur. München 1988, S. 279-288.
- Gunter SENFT: Völkerkunde und Linguistik. In: Themenheft Linguistik und Kulturanalyse. Zeitschrift für germanistische Linguistik 34.1/2 (2006), S. 87-104.

Jörg R. J. Schirra und Klaus Sachs-Hombach

### Anthropologie in der systematischen Bildwissenschaft: Auf der Spur des *homo pictor*

#### 1. Homo Sapiens – Homo Pictor

L'espèce humaine existe sur la terre depuis plus de 4 millions d'années. [...] Avec l'apparition de l'Homo sapiens, une révolution s'est opérée dans le mécanisme de la logique, dans le mode de pensée, dans la capacité d'abstraction et de synthèse. D'après nos connaissances actuelles, cette révolution n'a d'équivalent ni dans les précédentes étapes de l'histoire humaine ni dans aucune autre espèce animale. Le langage visuel, la capacité et le besoin de chercher un sens et un ordre préétabli dans les formes et les phénomènes de la nature, la recherche d'une communication avec des entités et des énergies immatérielles ou invisibles sont autant d'expressions de ces nouvelles acquisitions nées de cette révolution.<sup>1</sup>

Es hat Tradition, den Menschen als sprachbegabtes Tier zu charakterisieren. Dabei ist speziell an den Gebrauch prädikativer Satzstrukturen gedacht. Aber auch die merkwürdige Fähigkeit, Bilder zu verwenden, ist, nach allem was wir empirisch wissen, nur dem Menschen eigen. Gibt es begriffliche Gründe für diese empirische Koinzidenz? Ist, anders gefragt, der *homo sapiens* ganz wesentlich ein *homo pictor*? Eine solche Frage gehört ebenso gut in die Philosophische Anthropologie, wie sie der systematischen Bildwissenschaft eigenet. Von systematischer Bildwissenschaft sollte dann die Rede sein, wenn sich das intellektuelle Interesse der Frage zuwendet, was es grundsätzlich bedeutet, mit Bildern (als solchen) umgehen zu können.<sup>2</sup> Demnach stehen nicht einzelne Bilder im unmittelbaren Fokus des Interesses, sondern vielmehr die Fähigkeit, Bilder verwenden (erzeugen und rezipieren) zu können. Von Bildwissenschaftlern müssen folglich als Objekte ihrer Studien die Wesen, die über diese Eigenschaft verfügen, betrachtet werden. Noch präziser formuliert: Es geht um den *Begriff*, den wir uns von Wesen mit der erwähnten Fähigkeit auf sinnvolle und rational kontrollierte Weise bilden können (und sollten).

<sup>1</sup> Emanuel ANATI: La religion des origines. Paris 1999, S. 89.

<sup>2</sup> Klaus SACHS-HOMBACH: Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft. Köln 2003; Hans BELTING: Bild-Anthropologie. Entwürfe für eine Bildwissenschaft. München 2001; Horst BREDEKAMP: Bildwissenschaft. In: Metzler Lexikon Kunstwissenschaft. Stuttgart / Weimar 2003.

Diese grundsätzliche Fragestellung hat bereits Hans Jonas aufgeworfen. Unter dem Titel „Die Freiheit des Bildens – *Homo pictor* und die *differentia* des Menschen“ [1961] reflektiert er aus phänomenologischer Perspektive über den Stellenwert des Bildvermögens für den Begriff des Menschen. Jonas beginnt seine Überlegungen mit einem Gedankenexperiment zur Frage „Was ist Menschsein?“: Wie könnten, so fragt er, Weltraumforscher erkennen, ob es sich bei Wesen, denen sie auf einer anderen Welt begegnen, um „Menschen“ handelt? Natürlich ist vom Menschen hier nicht im Sinne der biologischen Gattung die Rede, ebenso wenig, wie ja die Philosophische Anthropologie mit der empirischen Anthropologie zusammenfällt.<sup>3</sup> An welche Symptome lassen sich also Kriterien knüpfen, die, der Charakterisierung *sapiens* entsprechend, bei jenen Wesen auf ‚Verstehen‘, ‚Geist‘, ‚Kultur‘, ‚Zivilisation‘ etc. schließen lassen? Das Vermögen, Bilder zu verwenden, mutet hierzu als eine besonders günstige Wahl an, da diese Fähigkeit einfacher als etwa das Sprachvermögen zu sein scheint, andererseits aber auch keine graduellen Übergänge zu rein biologisch erklärbaren Phänomenen erkennbar sind, wie sie etwa beim Werkzeuggebrauch auftreten. Würden die Astronauten aus dem Gedankenexperiment in einer Höhle künstlich erzeugte Linien und sonstige Farbkonfigurationen finden, d.h. Artefakte, die sie als Bilder interpretieren, dann wäre, so Jonas, ihre spontane Folgerung, dass es Menschen (im weiten Sinn) waren, die diese Artefakte gemacht haben. Was aber heißt es, Bilder verwenden<sup>4</sup> zu können? Was ist dazu vorausgesetzt?

Obwohl die Frage nach der Bildfähigkeit nicht auf einzelne Bilder abzielt, mag der Blick auf konkrete Exempel die Betrachtung auf wichtige Aspekte fokussieren. Nehmen wir etwa das in Abbildung 1 wiedergegebene Beispiel: Es handelt sich – um es zunächst möglichst neutral auszudrücken – um das Produkt der Beschäftigung eines Menschenaffen mit Papier und Stift, von dem Gardner & Gardner [1980] im Rahmen ihrer Untersuchung zur Sprachfähigkeit von Schimpansen berichtet haben. Ein Individuum namens Moja, das trainiert worden war, mit den Forschern mittels einer Teilmenge der *American Sign Language* zu kommunizieren, habe gelegentlich diese Linien spontan zu Papier gebracht. Ein Forschungsassistent, der direkt danach die Zeichen für ‚Was ist das?‘ an Moja richtete, bekam als Reaktion die Geste für ‚Vogel‘. Der Assistent bekundete, er habe diese Reaktion angesichts der Markierungen für plausibel gehalten und so verstanden, dass Moja damit angab, das Bild eines Vogels gemacht zu haben. Tatsächlich haben wir es hier mit einer empirischen Variante von Jonas’ Gedankenexperiment zu tun: Statt der unbekanntem We-

3 Vgl. etwa Ernst TUGENDHAT: Anthropologie als „erste Philosophie“. In: DERS.: Anthropologie statt Metaphysik. München 2007, S. 34–54.

4 Mit „Bildverwendung“ im hier gemeinten Sinn seien Bildproduktion und Bildrezeption zusammengefasst.

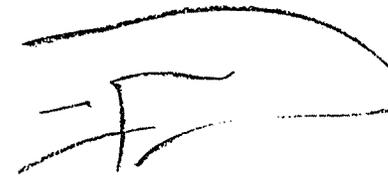


Abb. 1: Mojas Markierungen. Aus: Richard A. Gardner, Beatrice T. Gardner: Comparative Psychology and Language Acquisition. In: Thomas Sebeok, Donna Jean Umiker-Sebeok (Hgg.): Speaking of Apes. New York 1980, S. 287-330.

sen auf einem fremden Planeten sind wir mit einem Individuum einer anderen Erd-Spezies konfrontiert, der wir gemeinhin nicht zuschreiben, auch die spezifische Differenz des Menschseins aufzuweisen (was sich auch an dem Fehlen eines entsprechenden rechtlichen Status sehr deutlich zeigt); ein Individuum, das nun allerdings nicht nur bei der Herstellung des Artefakts beobachtet wurde, sondern von dem sogar einige weitere Umgangsformen damit bzw. ein Verhalten dazu (die Interaktion mit dem Assistenten) bekannt sind. Hat Moja hier also wirklich ein Bild verwendet? Und folgt daraus mit Notwendigkeit, wie Jonas annehmen müsste, dass wir Moja als *sapiens*, als Menschen im hier betrachteten Sinn behandeln sollten (mit allen zugehörigen – letztlich auch juristischen – Konsequenzen)?

Im Folgenden wird es darum gehen, näher zu bestimmen, wodurch sich Bildgebrauch auszeichnet, d.h. mit welchen Begriffen wir dabei eigentlich operieren (Abschnitte 4 und 5). Das führt uns zu zwei zentralen Begriffen: dem der Kontextbildung (7), der insbesondere eine Brücke zur Verwendung von Sprache schlägt; und dem der Ähnlichkeit (6), der sich als weit weniger trivial zeigt, als oft angenommen wird. Zuvor versuchen wir als methodologische Vorbereitung der begrifflichen Betrachtung kurz in zwei Abschnitten zu rekapitulieren, was gegenwärtig überhaupt mit begrifflichen Klärungen und mit deren Begründungen gemeint ist (2 und 3). Die anthropologische Dimension des Bildvermögens tritt dann in den Abschnitten 8 und 9 deutlicher hervor: Ausgehend von einem begriffsgenetischen Schema für den Bildbegriff werden Argumente für eine enge Wechselbeziehung zwischen Bildvermögen und Sprachvermögen aufgezeigt, die zwar der Grundannahme von Jonas zumindest teilweise widersprechen, doch seine Folgerungen für die Verknüpfung des unter dem Ausdruck *homo sapiens* gefassten (nicht biologisch verstandenen) Begriffs des Menschen und jenes mit *homo pictor* bezeichneten Begriffs verstärken.

## 2. Die Bestimmung des Menschen und der *linguistic turn*

Wenn im Folgenden von Sprache die Rede ist, dann ist vor allem an den Gebrauch prädikativer Satzstrukturen gedacht, also an Aussagen und davon abgeleitete Äußerungsarten (etwa entsprechende Frage- oder Imperativformen) mit ihrer typischen Binnengliederung, die weiter unten etwas genauer in den Blick genommen wird. In der Tat spielt das Sprachvermögen für die bildwissenschaftliche Betrachtung in jedem Fall eine überaus wichtige Rolle. Abgesehen davon, dass wir uns zur Vermittlung bildwissenschaftlicher Untersuchungen – notwendiger Weise, wie es scheint – der Sprache als zentralem Medium bedienen, erfüllt das Sprachvermögen auch den Zweck eines Referenzsystems für die anthropologische Bestimmung: Nur wenn sich aus der jeweils untersuchten Begriffserläuterung auch die Charakterisierung des Menschen als sprachbegabtes Tier ableiten lässt, kann davon die Rede sein, dass die spezifische Differenz des Menschseins erfasst worden ist. Sprachvermögen kann also in bildanthropologischen Betrachtungen auch inhaltlich – nämlich als Kriterium des Erfolgs – auftreten.

Über die mediale und die inhaltliche Bedeutung hinaus spielt Sprache aber auch methodologisch eine besondere Rolle, geht es bei den hier erörterten Fragen doch nicht um empirische (d.h. vorgegebene Begriffe lediglich verwendende) Untersuchungen – also etwa zum Vorliegen oder Nicht-Vorliegen gewisser biologischer, psychologischer oder soziologischer Zusammenhänge –, sondern um das Klären begrifflicher Beziehungen selbst. Nun soll im Weiteren, der modernen philosophischen Argumentationstheorie<sup>5</sup> folgend, unter einem Begriff die *interindividuell verfügbare Instanz zur Überprüfung der Geltung von prädikativen Sätzen* verstanden werden. Das ist recht abstrakt formuliert, meint letztlich aber nichts anderes, als dass ein Begriff einer Unterscheidungsgewohnheit entspricht, die sozial kontrolliert wird (also prinzipiell intersubjektiv verfügbar wie auch veränderbar bleibt), und deren Anwendung auf eine konkrete Situation zu empirischem Wissen führt (oder zumindest führen kann). Häufig wird daher ein Begriff auch gleichgesetzt mit der Bedeutung eines entsprechenden prädikativ gebrauchten Terminus, wie etwa ‚Mensch sein‘.

Die Klärung begrifflicher Beziehungen betrifft mithin Zusammenhänge zwischen solchen speziellen Unterscheidungsgewohnheiten. Im Verlauf der Philosophiegeschichte hat sich allerdings die Auffassung davon, was Begriffe sind und wie sie sich klären lassen, mehrfach gewandelt. Die letzte und hier bedeutsame Wandlung wird häufig unter der Bezeichnung *linguistic turn* abgehandelt. Insbesondere verändert sich dabei das Verhältnis zwischen Begriff-

5 Arno Ros: Begründung und Begriff. Hamburg 1989/1990.

fen und Sprachvermögen. War zuvor angenommen worden, Begriffe seien 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 wäre und der Gebrauch von Begriffen (das Denken) auch ganz ohne Sprachvermögen möglich bliebe, so geht man nun davon aus, dass sich der Begriff des Begriffs nur dann sinnvoll konzipieren lässt, wenn man Begriffe als etwas auffasst, was prinzipiell sprachlich (genauer: kommunikativ) konstituiert wird.

Der *linguistic turn* stellt daher eine radikale Abkehr von der bewusstseinstheoretischen Position dar. Er ist letztlich eine Folge der Unmöglichkeit, solipsistische Konsequenzen zu umgehen, wenn mit den Begriffen gleichsam die Prüfkriterien für die Geltung prädikativer Äußerungen in einem – jeweils von anderen unzugänglich gedachten – Einzelbewusstsein eingeschlossen bleiben und so ihre eigentliche kommunikative Funktion gar nicht erfüllen können. Diese fatale Konsequenz der im wesentlichen von den Philosophen des 17. bis 19. Jahrhunderts ausgearbeiteten bewusstseinstheoretischen Position war zu Beginn des 20. Jahrhunderts ausgesprochen virulent geworden und führte insbesondere zu zwei unterschiedlichen Reaktionen. In phänomenologischen Betrachtungen wurde versucht, das Problem mithilfe so genannter „Wesenschauen“ zu lösen – letztlich ein Rückgriff auf die Position der antiken Philosophie, in der die Prüfinstanzen (etwa in Form der platonischen Ideen) als außerhalb von Raum und Zeit befindliche unabhängige Wesenheiten konzipiert waren. Der je individuell verwendete Begriff sollte dann an der jeweiligen, irgendwie überindividuell vorgegebenen Wesenheit – dem *Eidos* – ausgerichtet werden und so das Problem der solipsistischen Abkapselung umgehen.

Alternativ wurde versucht, die Begriffe für psychische Phänomene (zu denen in der bewusstseinstheoretischen Philosophie auch die Begriffe selber gehörten) anders zu fassen. Psychologische Ausdrücke werden nicht länger als etwas verstanden, das auf völlig private Entitäten im Bewusstsein verweist, sondern als sprachlicher Ausdruck von bestimmten, durchaus intersubjektiv zugänglichen Verhaltensmerkmalen. Wurden z.B. gewisse mentale Phänomene in der bewusstseinstheoretischen Philosophie aufgefasst als etwas, das Handlungen kausal auslöst und als solches von der jeweils ausgelösten Handlung selbst unabhängig sein musste, so werden diese Phänomene nun als (durchaus komplexe) *Aspekte* jener Handlungen betrachtet und damit als davon begrifflich gerade nicht unabhängig. Der Witz von Begriffen, die durch Ausdrücke wie ‚Vorstellungen‘, ‚Gedanken‘ und eben auch ‚Begriffe‘ artikuliert werden, ist dann in bestimmten, im Einzelnen jeweils genauer festzulegenden Betrachtungsweisen der entsprechenden Verhaltenskomplexe gegeben. Insbesondere ist Sprechen bzw. Kommunizieren eine Art solcher Handlungen, die sich für Begriffe als konstitutiv ergibt. Der *linguistic turn* ist mithin eine direkte Kon-

sequenz dieser handlungstheoretischen Umdeutung der Begriffe für psychische Phänomene.

Nun folgt aber unmittelbar: Wenn Begriffe wesentlich sprachlich vermittelt und bestimmt sind, dann hängt alles, was am Gebrauch von Begriffen hängt, davon ab, dass das entsprechende Sprachvermögen vorhanden ist. Insofern führt der *linguistic turn* (bzw. die handlungstheoretische Umdeutung psychischer Phänomene) auch dazu, dass Sprachvermögen nicht nur irgendein Symptom des Menschseins (im anthropologischen Sinn) ist, sondern sein ganz zentrales Kriterium.<sup>6</sup> Wo im Vergleich zu diesem Kriterium steht aber das Bildvermögen?

### 3. Argumentationstheoretische Basis der Debatte

Eine begriffliche Klärung des Bildvermögens und seines anthropologischen Status läuft letztlich darauf hinaus zu analysieren, worin die ‚logisch-begrifflichen Voraussetzungen (oder mit Kants Worten: die „transzendentalen Bedingungen der Möglichkeit“) bestehen, die wir jeweils unterstellen müssen, wenn wir von Wesen mit der Fähigkeit zum Bildgebrauch (bzw. zum Sprachgebrauch im Vergleich) reden. Und mehr noch: sie sollte untersuchen, wie wir diese Voraussetzungen korrekt begründen können.

Zu diesem Zweck sei zunächst an einige methodologische Erkenntnisse der philosophischen Argumentationstheorie erinnert. Sie unterscheidet zwischen Begriffserläuterungen und Begründungsversuchen für Begriffserläuterungen.<sup>7</sup> Knapp gefasst, führt eine Begriffserläuterung in einem Dissens dazu, dass ein strittiger Begriff – d.h. eine interindividuell kontrollierte Unterscheidungsfähigkeit – bestimmt wird durch eine logische Kombination anderer Begriffe für den gleichen Phänomenbereich. Definitionen sind typische Beispiele dieses Verfahrens. Nun werden allerdings jene zur Erläuterung verwendeten Begriffe bei weiterer Rückfrage oft wiederum selbst unter anderem mit dem zuerst erläuterten Begriff erklärt. Systeme aus sich wechselseitig bestimmenden Begriffen nennt man auch ein *Begriffsfeld*. Solche zyklischen Abhängigkeiten lassen sich meist lösen, indem eine kleine Menge von Grundbegriffen eines Begriffsfel-

<sup>6</sup> Deshalb ist die oben angesprochene Vermutung, Sprache sei als Medium zur Vermittlung bildwissenschaftlicher Anstrengungen unumgänglich, in der Tat korrekt: Die mediale Bedeutung von Sprache für die Bildwissenschaft ergibt sich direkt aus ihrer methodologischen Relevanz. Nochmals sei ausdrücklich darauf hingewiesen, dass hier nicht die phonetische oder graphische Form unserer Sprachen entscheidend ist, sondern das Vorhandensein einer Kommunikationsform mit der typischen logisch-funktionalen Gliederung der Aussagen (s.u.).

<sup>7</sup> Arno Ros: Was ist Philosophie? In: Richard RAATZSCH (Hg.): Philosophieren über Philosophie. Leipzig 1999, S. 36–58.

des festgelegt wird, auf die jeder andere Begriff des Systems logisch zurückgeführt werden kann. Die Grundbegriffe selbst sind dann das zunächst nicht weiter hinterfragbare Ende der Erklärungskette. Ein Dissens um einen Begriff lässt sich also beheben, sofern die Grundbegriffe des betrachteten Begriffsfeldes von allen Parteien als unbedenklich akzeptiert werden.

Ist das nicht der Fall, muss es darum gehen, Gründe anzugeben, die von allen Beteiligten akzeptiert werden können, für oder wider den einen oder anderen Aspekt am System der Grundbegriffe. Im Gegensatz zu den stets *intern* zu dem Begriffsfeld verlaufenden Begriffserläuterungen müssen diese Gründe jedoch *begriffsfeld-extern* sein. Denn es ist offensichtlich, dass eine Begründung von Grundbegriffen – entsprechend den Axiomen einer Theorie – sich nicht durch eine ‚logische‘ Ableitung innerhalb des Begriffsfeldes ergeben kann. An dieser Stelle treten *begriffsgenetische* Betrachtungen ins Spiel, d.h. der Vorschlag, das fragliche Begriffsfeld bzw. das System von Grundbegriffen zu betrachten als eingeführt durch eine systematische Kombination von zunächst voneinander unabhängigen und von allen Parteien akzeptierten Begriffsfeldern, die üblicherweise eine gegenüber dem kombinierten Feld einfachere innere Struktur aufweisen. Während die Instanzen von Gegenständen, die unter die einfacheren Begriffsfelder fallen, Eigenschaften aus dem jeweils anderen Begriffsfeld höchstens in kontingenter Weise aufweisen, kommen unter dem kombinierten Begriffsfeld Instanzen vor, die Eigenschaften aus beiden elementareren Begriffsfeldern auf systematische Weise gekoppelt zeigen. Dabei legt das Schema der Begriffsfeldkombination zusammen mit den internen Regeln der elementareren Felder diejenigen Regeln fest, die für begriffsfeld-interne (logische) Begriffserläuterungen im kombinierten Feld gelten: Damit ist eine Begründung der Axiome selbst gegeben. Die begriffsgenetische Betrachtung baut darauf, dass sich zweierlei zeigen lässt: (1) In keinem der elementareren Felder lassen sich bestimmte, für uns relevante Begriffe definieren. (2) In dem kombinierten Feld ist es möglich, jene Begriffe zu definieren. Es gibt damit eine zusätzliche Möglichkeit, einen Dissens über diese Begriffe und ihre Eigenheiten zu lösen, sofern die Parteien die elementareren Begriffsfelder akzeptieren, das Schema der Kombination annehmen und überhaupt Interesse daran haben, gemeinsame Begriffe der betrachteten Art zu etablieren.

Es liegt nahe, solche begriffsgenetischen Betrachtungen als wesentliche Bestandteile auch der Philosophischen Anthropologie und der Bildwissenschaft zu verstehen: Die Darstellung einer möglichen begriffsgenetischen Ableitung eines Begriffsfeldes, das einen vorgeschlagenen Begriff des Menschen enthält, aus weniger komplexen Begriffsfeldern, d.h. vor allem Begriffsfeldern mit den Begriffen für einfachere Kreaturen als den Menschen, müsste gerade zeigen, auf welchen – ‚transzendentalen‘ – Voraussetzungen (d.h. welchen Bedin-

gungen der Möglichkeit) Sprachvermögen und Bildvermögen beruhen und wie beides miteinander und mit den anderen *differentia* des Menschen zusammenhängt.

Es sei nochmals ausdrücklich darauf hingewiesen, dass solche begriffsgenetischen Betrachtungen weder mit begriffsgeschichtlichen Überlegungen (in diesem Fall zum Begriff ‚Mensch‘) noch mit entwicklungsbiologischen Argumentationen (zur Tiergattung *homo sapiens*) oder kulturhistorischen Erwägungen (etwa zum Bildermachen) identisch sind: Die drei letztgenannten sind im wesentlichen *empirische* Untersuchungen, die sich im ersten Fall auf die konkrete Abfolge verschiedener Versionen und Vorversionen des betrachteten Begriffes in einer bestimmten Kultur richten, im anderen Fall die tatsächlichen Übergangsschritte der biologischen Gattungen hin zum Menschen im Auge haben, und im dritten Fall auf den wirklichen Hergang beim Entstehen gewisser Kulturtechniken in einem bestimmten Kulturkreis abzielen. Im Gegensatz dazu verfolgt die begriffsgenetische Betrachtung das Ziel, nicht empirische sondern *begrifflich-strukturelle* Zusammenhänge zwischen Begriffen zu untersuchen und so zur Gestaltung der Unterscheidungsgewohnheiten beizutragen, die von uns im Weiteren zur Gewinnung empirischer Kenntnisse sanktioniert sind.

#### 4. Der Bildbegriff bei Hans Jonas

Im oben erwähnten homo pictor-Aufsatz hat Hans Jonas mehrere Kriterien zusammengestellt, die für das Zuschreiben des Prädikats ‚Bild‘ wichtig sind und im Weiteren – wenn auch teilweise in etwas veränderter Form – im Fokus der Aufmerksamkeit bleiben. Insbesondere sind, nach Jonas, Gegenstände dann Bilder, wenn sie sich durch die folgenden Eigenschaften auszeichnen:<sup>8</sup>

1. *Ähnlichkeit* zu einem anderen Ding: Diese Relation zwischen zwei Gegenständen nennt Jonas an erster Stelle. Die Ähnlichkeit mag sich spontan zeigen oder erst „auf Wunsch“ (als eine nur vorgestellte oder anders vermittelte Ähnlichkeit) erkennbar werden.
2. *Interne Intentionalität*: Nur eine absichtlich hervorgebrachte Ähnlichkeit soll zählen. Damit wird die eigentlich symmetrische Ähnlichkeitsrelation unsymmetrisch, denn etwas Abgebildetes ist seinem Bild zwar ähnlich, soll aber nicht wiederum umgekehrt als dessen Bild gelten. Diese Absicht müsse am Bild erkennbar sein. Daher soll nach Jonas die ursprüngliche, an

<sup>8</sup> Die hier gegebene Darstellung stimmt dabei insgesamt zwar sinngemäß, aber weder in der genauen Reihenfolge noch in einigen der verwendeten Termini mit Jonas' Original überein: Die Abweichungen dienen der größeren Übersichtlichkeit und besseren Anschlussfähigkeit an die folgende Diskussion.

den Bildhersteller gebundene Absicht als eine „innere“ Intentionalität selbstständig im Resultat des Herstellens (dem Bild) erscheinen. Dort teile sie sich dem Betrachter spontan als Intentionalität der Darstellung mit.

3. *Repräsentationalität*: Der Gegenstand, der als Bild aufgefasst wird, vertritt in gewisser Weise etwas anderes – stellt es dar. Diese Eigenheit ist in den drei folgenden Punkten näher bestimmt:
  - a. *Ungleichartigkeit*: Das Bild mag dem abgebildeten Gegenstand zwar ähnlich sein, aber es ist keinesfalls ein Exemplar exakt der gleichen Art: Es muss eine gewollte „Unähnlichkeit in der Ähnlichkeit“ bestehen. Damit steht das Bild im Gegensatz zur Täuschung, die ungewollt unterläuft. Bildliche Ähnlichkeit beschränkt sich auf die Oberfläche. Diesen Aspekt des Unähnlichen versteht Jonas als konstitutiv für das Bildsein und nennt ihn „ontologische Unvollständigkeit“.
  - b. *Grade der (Un-)Ähnlichkeit*: Welche der oberflächlichen Aspekte ähnlich und welche anderen unähnlich sind, ist dabei vor allem eine Wahl des Bildverwenders: Es handelt sich um Freiheitsgrade, die im Sinne von Ökonomie und Expressivität genutzt werden können.
  - c. *Emanzipation von Wörtlichkeit*: Diese Freiheitsgrade erlauben dem Bildverwender, gezielt Abweichungen zur Verbesserung der Repräsentationalität einzusetzen. Konventionalisierungen entsprechender Abstraktionen und Stilisierungen können dabei nach Jonas nicht nur zur Bildung von Stilkanons führen; sie öffnen konzeptuell den Weg bis zu ideographischen Schriften.
4. *Visualität*: Zumindest im hier betrachteten engeren Sinn ist von Bildern nur dann die Rede, wenn wesentlich der (menschliche) Gesichtssinn angesprochen ist. Dabei soll das Sehen sich, nach Jonas, dadurch unter unseren Sinnen auszeichnen, dass hier, wie sonst nirgends, ein Ding in vielen verschiedenen Erscheinungen doch als dasselbe Ding (mit einer bestimmten ihm eigenen Form) auftreten kann: ein Verhältnis, das gewissermaßen die Relation *Eidos* zu Einzelding paradigmatisch vorführe und so zum Ausgangspunkt weiterer Abstraktionsschritte werden könne. Ebenso können unterschiedliche Gegenstände in gleicher visueller Gestalt auftreten und also durch dieselben Bilder abgebildet sein. So werde in Bildern Allgemeinheit sinnfällig.
5. *Herstellen von Distanz*: Mit einem Bild werden auf mehrfache Weise Distanzierungsschritte wirksam, die letztlich in der Unterscheidung ‚Darstellung‘ (Bild), ‚Darstellendes‘ (Bildträger), ‚Dargestelltes‘ (Bildinhalt) als drei im Bildbegriff zu berücksichtigenden ontologischen Schichten münden. Damit bildet die Distanzierung zugleich die Grundlage der Repräsentationalität.

- a. *Inaktivität*: Ein Aspekt der Distanzierung besteht darin, dass das Dargestellte im Bild zwar gegenwärtig zu sein scheint, aber dem normalen „Kausalverkehr der Dinge“ entzogen bleibt. Entsprechend mag zwar eine Bewegung dargestellt sein, aber deren normale Folgen – der Ortswechsel und weitere dadurch verursachte Bewegungen – treten nicht ein: Die Bewegung bleibt in „statische Gegenwart“ gebannt.
- b. *Selbstverleugnung des Bildsubstrats* (Bildträger): Im Bild wird das Darstellende – der Bildträger – zugunsten des Abgebildeten ignoriert. Nach Jonas taucht damit auch die Vorgeschichte dieses Gegenstands, zu der natürlich auch der Herstellungsprozess gehört, unter.

## 5. Zeichen, wahrnehmungsnahe Zeichen, Bilder

Die bisher nachgezeichneten Überlegungen von Jonas könnten dazu verleiten anzunehmen, dass Bilder unabhängig von ihrer Verwendung existierten und erst in einem zweiten Schritt von Wesen mit passenden Eigenschaften als Bilder gebraucht würden. Das ist natürlich Unsinn: Von einem Bild kann nur die Rede sein, insofern es als Bild *verwendet*, d.h. als solches hergestellt oder rezipiert wird. Daher sollte sich der Bildbegriff direkt auf die zugrunde liegende(n) Gebrauchssituation(en) beziehen. In diesem Zusammenhang fällt auf, dass drei von Jonas' Kriterien – *Herstellen von Distanz*, *Intentionalität* und *Repräsentationalität* – ebenfalls ganz charakteristisch sind für die Situation eines (recht allgemein bestimmten) Zeichengebrauchs: Im Zentrum einer solchen Situation steht ein Gegenstand, mit dessen Hilfe ein Sender einem Empfänger – als zwei weiteren notwendigen Bestandteilen der Situation – etwas in der Regel nicht zugleich Anwesendes vergegenwärtigen möchte. Daher kann der Gebrauch von Bildern, in dem allein sich ihre Bildhaftigkeit manifestiert, durchaus als eine Art Interaktion vom Typus *Zeichenhandlung* betrachtet werden.

Es mag irritierend wirken, Bilder als eine besondere Art von Zeichen zu bestimmen, scheint es doch ein zentraler Verwendungszusammenhang von Bildern zu sein, dass eine Person alleine und ganz für sich ein Bild betrachtet. Hingegen wird der Verwendung von Zeichen eine prinzipiell dialogische Situation zugrunde gelegt: Jemand gibt jemand anderem etwas zu verstehen (oder intendiert das mit der Zeichenhandlung doch zumindest). Da speziell sprachliche Zeichen notwendigerweise nicht einfach in der Natur vorzufinden sind, sondern von einem Sprecher produziert sein müssen, ergibt sich sofort, dass ihr Gebrauch wesentlich eine Interaktion zwischen zwei Partnern – also eine Kommunikation – ist. Bei Bildern gibt es demgegenüber, jedenfalls auf den

ersten Blick, durchaus Fälle, die natürlich, also unabhängig von einem Sender, entstanden zu sein scheinen: Man denke etwa an spiegelnde Flächen.

Mehr als dem Gespräch ähnelt das einsame Betrachten eines Bildes dann vielleicht dem Lesen, das ja auch in der Regel Einer, der möglichst dabei nicht gestört werden möchte, für sich alleine betreibt. Doch handelt es sich, genau besehen, auch beim Lesen um eine kommunikative Aktivität, bei der allerdings einer oder mehrere der Interaktionspartner internalisiert sind und nur in der Vorstellung als Gegenüber auftreten. Ist es nun aber nicht so, dass ich, wenn ich ein Spiegelbild betrachte, ganz analog *mir selbst* (in der Rolle eines anderen, der von jener Stelle aus schauen könnte) etwas *zeige*? Und mache ich nicht *mich selbst* (als ein innerlich gespielter Anderer), wenn ich in einer Galerie ein Bild studiere, mit dem Bild auf dieses oder jenes aufmerksam? Es ist also durchaus sehr plausibel, auf diese Weise auch einsame Bildverwendungen auf kommunikative Handlungen zurückzuführen: Es gäbe demnach, wo immer von Bildern die Rede ist, notwendig einen (gegebenenfalls interisorierten und vielleicht sogar nur imaginierten) Sender und einen (ebensolchen) Empfänger, die auf je spezifische Weise mit Hilfe eines (Bild-) Zeichens kommunikativ miteinander interagieren (zumindest in der Vorstellung). Genau in diesem noch recht unspezifischen Sinn sprechen wir im Folgenden von einer *Zeichenhandlung*.<sup>9</sup>

Wird nun Zeichengebrauch als Oberbegriff verwendet, dann könnte die spezifische Differenz des Bildgebrauchs anderen Zeichenhandlungsarten gegenüber zu tun haben mit den beiden übrigen Kriterien nach Jonas. In der Tat führt zunächst *Ähnlichkeit* – als konstitutive Größe des Zeichengebrauchs betrachtet – zum Begriff der *wahrnehmungsnahen Zeichen*. Darunter sollen die Zeichen fallen, bei deren Gebrauch die Zeichenverwender wesentlich auf diejenigen Wahrnehmungskompetenzen zurückgreifen, die sie eigentlich einsetzen, wenn sie das Dargestellte selbst (d.h. ohne Vermittlung durch ein Zeichen) wahrnehmen. Die Klasse der wahrnehmungsnahen Zeichen kann dann weiter gemäß verschiedener Sinnesmodalitäten aufgespalten werden. Damit ist bereits der Weg, den Begriff der Ähnlichkeit handlungstheoretisch zu bestimmen, angedeutet: ein Weg, der letztlich erläutert, wie Ähnlichkeit in die Gebrauchssituation der Zeichenhandlung eingebettet wird.

<sup>9</sup> Übrigens lässt sich damit auch die etwas eigentümliche Konzeption der *inneren Intentionalität* bei Jonas besser fassen: Dass das Bild selbst die originale Intentionalität des Herstellers zu übernehmen scheint, kann man auch dadurch erläutern, dass jeder, der den Bildträger als Bild rezipiert, prinzipiell in der Lage ist, sich einen entsprechenden Sender (Bildhersteller) vorzustellen und dessen unterstellte Herstellungstentionalität als innere Intentionalität des Bildes zu verwenden. Für die spontane Interpretation der Raumfahrer in der Höhle nach Jonas' oben erwähntem Gedankenexperiment führt diese Auffassung dann allerdings direkt zu einem ernstem Problem.

Es lohnt sich, an dieser Stelle zunächst kurz zu überdenken, wovon hier, wie schon bei Jonas, *nicht* die Rede sein soll: Denn den Ausdruck ‚Bild‘ benutzen wir in vielfältiger Weise, und nicht alle diese Verwendungsweisen sind gleichermaßen geeignet, die mit dem Ausdruck ‚Bild‘ unmittelbar artikulierten Unterscheidungsgewohnheiten deutlich werden zu lassen. Ist etwa von Weltbildern oder Feindbildern oder gar Weibsbildern die Rede, so geht es wohl nicht tatsächlich um Bilder in dem hier eigentlich betrachteten Sinn; vielmehr wird dabei der Bildbegriff in einer *uneigentlichen* Weise verwendet. Daher sollten metaphorische Übertragungen von direkten Anwendungen des betrachteten Bildbegriffs unterschieden werden. Das gilt übrigens – jedenfalls auf den ersten Blick – auch für die so genannten ‚inneren‘ (oder ‚mentalen‘) Bilder, ist für sie doch weder klar, in welchem Sinne sie eigentlich Teil einer Zeichenhandlung sein sollen, noch, auf welche Weise bei ihnen die Wahrnehmungskompetenzen für das scheinbar Dargestellte zum Tragen kommen könnten.

Abgesehen von metaphorischen Verwendungen muss damit gerechnet werden, dass neben den zentralen Instanzen auch periphere Bildfälle vorkommen: Bei den ersten sollte der gewählte Bildbegriff voll anwendbar sein, während bei letzteren seine charakteristischen Eigenschaften im Sinne der Wittgensteinschen Familienähnlichkeit mehr oder weniger aufgeweicht in Erscheinung treten. So kann bei Strukturbildern – etwa graphische Darstellungen zu Wahlwanderbewegungen – die Ähnlichkeitsbeziehung nicht ohne weiteres mit Blick auf entsprechende Wahrnehmungskompetenzen bestimmt sein: Wahlwanderbewegungen werden nun einmal nicht visuell wahrgenommen. Wie Jonas gehen wir gemeinhin davon aus, dass darstellende Bilder – d.h. Bilder von raumzeitlichen Konfigurationen materieller Gegenstände – zentraler seien als Strukturbilder.<sup>10</sup> Dabei kommt den dargestellten Gegenständen selbst bereits eine Besonderheit zu, die für das Bildvermögen aus handlungstheoretischer Perspektive von großer Bedeutung ist, hängt sie doch damit zusammen, dass wir uns vom Hier und Jetzt distanzieren können: Diese Gegenstände sind *individuiert*. In der Philosophie spricht man auch von *sortalen* Gegenständen.

<sup>10</sup> Der Terminus ‚darstellendes Bild‘ ist sicher unglücklich gewählt, stellen doch auch Strukturbilder durchaus etwas dar. Er hat sich aber in Ermangelung einer plausiblen Alternative etabliert. Angemerkt sei, dass darstellende Bilder weder notwendig ‚naturalistisch‘ (also dem tatsächlichen visuellen Erscheinungsbild möglichst ähnlich) sein müssen (vgl. z.B. Kupferstiche), noch sind sie auf Abbildungen im engeren Sinn (also von einer tatsächlich bestehenden Situation) beschränkt (Gegenbeispiel: Bilder fiktiver Szenen).

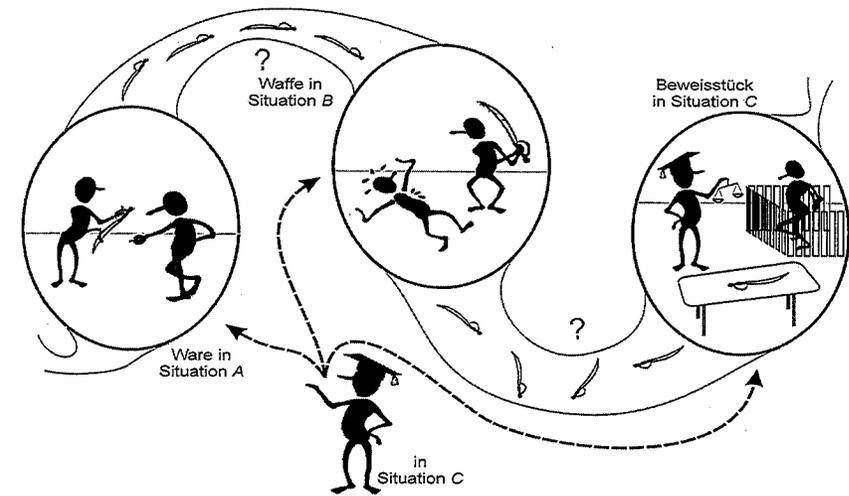


Abb. 2: Identität eines sortalen Gegenstands als Element verschiedener Kontexte

## 6. Kontextbildung und die Distanzierung vom Hier und Jetzt

Wir nehmen es üblicherweise als gegeben hin, dass unsere Welt (neben anderem) vor allem aus individuellen materiellen Gegenständen besteht: aus Teekannen und Stühlen, Äpfeln, Mäusen, Autos und Straßen etc. – Dinge, denen wir zu ganz verschiedenen Zeiten und teilweise auch an ganz unterschiedlichen Orten als *denselben* Dingen begegnen, selbst wenn sie sich in der Zwischenzeit verändert, bisweilen sogar sehr stark umgeformt haben (z.B. Raupe zu Schmetterling). Nur wer diese sortale Individuation beherrscht, kann durchschauen, ob es sich zu zwei verschiedenen Zeitpunkten um *denselben* Gegenstand handelt oder um zwei nur *gleichartige* Gegenstände, und entsprechend handeln. Um die Klärung einer solchen Identität geht es etwa, wenn ein Gericht festzustellen versucht, ob es sich bei dem Dolch, der jetzt auf dem Aservatentisch liegt, dem spitzen Gegenstand, der das Opfer vor einem Jahr am anderen Ende der Stadt niedergestreckt hat, und dem Messer, das der Angeklagte vor 13 Monaten in der Nachbarstadt gekauft hat, um ein und denselben Gegenstand handelt (vgl. Abb. 2). Auch die möglichen Transformationen, die einen Gegenstand der betrachteten Sorte im Verlauf seiner raumzeitlichen Lebenslinie verwandeln können, ohne seine Identität zu ändern (etwa das Rosten des Messers, dessen Verschmutzung oder das Umbiegen der Messerspitze bei unsachgemäßem Gebrauch, nicht aber das Austauschen der Klinge), sind nur als Teil des sortalen Begriffs determiniert.

Doch sollten wir uns klar machen, dass weder für sehr kleine Kinder noch für selbst relativ hoch entwickelte Tiere nachweisbar ist, dass sie in der Lage

sind, sich einem Objekt als etwas zu nähern, dem sie zu unterschiedlichen Zeiten als *ein und demselben* Individuum gegenüber treten könnten. Uns an das Umgehen mit sortalen Gegenständen Gewöhnten fällt es eigentümlich schwer, uns vorzustellen, was das Fehlen dieser Befähigung bedeutet.<sup>11</sup> Dinge werden dann überhaupt nur als etwas je konkret in der aktuellen Verhaltenssituation Anwesendes aufgefasst: Wendet sich die Aufmerksamkeit von ihnen ab, entschwinden sie dem Wesen unwiederbringlich. Eine weitere Begegnung führt höchstens dazu, nun einen *gleichartigen* Gegenstand wahrzunehmen. Wesen, die mit sortalen Gegenständen umgehen, können hingegen zwischen der je augenblicklichen Erscheinung eines sortalen Gegenstands, also wie man ihn in einer bestimmten Verhaltenssituation wahrnimmt, und dem Gegenstand als solchem, der im Prinzip die Gesamtgeschichte (mit vergangenen, zukünftigen oder auch potentiellen Erscheinungsweisen) umfasst, unterscheiden. Nur unter einem sortalen Gegenstandsbegriff werden daher die verschiedenen ‚Lebenslinien‘ (die Reaktionen) von einem, um ein literarisches Beispiel zu bemühen, echten und einem nur scheinbar echten Reh erkennbar: „*Gab ihm einen Stips; da war es aus Gips*“.<sup>12</sup>

Es ist wichtig, sich hier klarzumachen, dass sortale Gegenstände nie isoliert vorkommen: Es macht immer nur Sinn, von ihnen als etwas zu reden, was, da es prinzipiell in mehr als einem Verhaltenskontext existiert, als Figur vor einem jeweiligen, prinzipiell variablen Hintergrund in Erscheinung tritt. Mit ‚Kontext‘ soll hier einerseits jede endliche, strukturierte Menge von intentionalen sortalen Gegenständen bezeichnet werden: also eine Reihe von miteinander in Beziehung stehenden Gegenständen, insoweit jemand seine Aufmerksamkeit auf sie richtet (etwa durch Wahrnehmen). Allerdings können Kontexte ebenso gut als Verhaltenssituationen (für Wesen mit entsprechend komplexen Verhaltensweisen) charakterisiert werden. Es liegt zwar nahe, Situationen, auf die ein Verhalten bezogen ist, einfach als Gegenstandsmengen aufzufassen; doch sollte man sie vielmehr mit Wittgenstein als Systeme von Sachverhalten betrachten: das, was der Fall ist *dann und dort*. Denn jede ‚Verhaltenssituation‘ bildet ein Interpretationsangebot mit vielen Auslegungsmöglichkeiten.<sup>13</sup>

11 Bestimmte neurologische Pathologien scheinen gerade diese Fähigkeit auszuschalten, vgl. Alexander R. LURIA: *Der Mann, dessen Welt in Scherben ging*. Reinbek 1991; Oliver SACKS: *The Man Who Mistook His Wife for a Hat*. New York 1986, S. 220; sowie – in einer cineastischen Dramatisierung – den Spielfilm *Memento* von Christopher NOLAN [2000].

12 Frei nach Joachim RINGELNATZ: *Im Park*. In: *Reisebriefe eines Artisten* [1927].

13 Die Charakterisierung als eine Gegenstandsmenge ist gewissermaßen eine *figurale* Sicht auf einen Kontext. Ihr tritt eine *mediale*, d.h. figurale Gliederungen nur anbietende Sicht als interpretations-offene potentielle Verhaltenssituation zur Seite (vgl. Jörg R. J. SCHIRRA, Klaus SACHS-HOMBACH: *To Show and To Say. Comparing the Uses of Pictures and Language*. In: *Studies in Communication Sciences* 7/2 (2007),

Ein Kontext entspricht dem, was man umgangssprachlich oft auch als das ‚Hier und Jetzt‘ bezeichnet. Genau betrachtet kann es sich dabei nicht um einen einzelnen quasi ausdehnungslosen (physikalischen) Zeitpunkt handeln, denn die Aufmerksamkeitsspanne der Wahrnehmung überdeckt stets mehr oder weniger ausgedehnte Zeitintervalle. Affektive Bindungen auf einen (erwünschten/gefürchteten) Gegenstand erhalten sie zudem aufrecht: Auch wenn es gerade nicht wahrgenommen werden kann, bleibt das erwünschte/gefürchtete Objekt so Teil der aktuellen Verhaltenssituation. Kontexte kann man sich als zusammenhängende, mehr oder weniger ausgedehnte Raumzeit-Blasen um ein Wesen herum vorstellen: soweit die aktuelle Aufmerksamkeit jenes Lebewesens – sein „Merk-“ und „Wirknetz“ in den Worten von Uexkülls – eben reicht.<sup>14</sup> Damit bleibt für gewöhnlich die Aufmerksamkeit eines Lebewesens innerhalb seiner aktuellen Verhaltenssituation, also auf *einen* Kontext beschränkt.

Dass ein Lebewesen seine Aufmerksamkeit auf Kontexte selbst richtet – also auch mehrere Kontexte miteinander vergleichen kann – braucht nach bisherigem Wissen zur Erklärung tierischen Verhaltens nicht herangezogen zu werden. Neben den unmittelbaren Reaktionen auf die direkte Umgebung (und gegebenenfalls autonomen internen Steuerzyklen) genügen dazu Präge- und Habitationskonzepte, für deren Funktion keine Vergegenwärtigung einer anderen Situation nötig ist.<sup>15</sup> Für den Umgang mit sortalen Gegenständen spielt hingegen die Fähigkeit eine zentrale Rolle, die Aufmerksamkeit *frei* auf *beliebige* Kontexte richten zu können. Denn man muss sich vom jeweiligen Hier und Jetzt distanzieren und auch auf andere Gegebenheitsweisen eines Gegenstands orientieren können. Um im eben angeführten Bild zu bleiben: Es geht nicht nur darum, die Raumzeit-Blase der aktuellen Verhaltenssituation räum-

S. 35–62, Abschn. 4.2; siehe auch Ludwig WITTGENSTEIN: *Tractatus logico-philosophicus*. London 1922, §1).

14 Siehe Jakob von UEXKÜLL: *Umwelt und Innenwelt der Tiere*. Berlin 1909, Vgl. auch Jörg R. J. SCHIRRA: *Täuschung, Ähnlichkeit und Immersion. Die Vögel des Zeuxis*. In: Klaus SACHS-HOMBACH, Klaus REHKÄMPER (Hgg.): *Vom Realismus der Bilder. Interdisziplinäre Forschungen zur Semantik bildhafter Darstellungsformen*. Magdeburg 2000, S. 119–135.

15 Früher gemachte Erfahrungen mit einer Art von Gegenstand haben in dem Fall sozusagen das Gefüge von Merk- und Wirknetz selbst so verschoben, dass eine erneute Begegnung mit Exemplaren, die für das Wesen unter jene Gegenstandssorte fallen, generell zu veränderten Verhaltensweisen führt. Dass einfache Lebewesen, etwa bei der Prägung auf Elterntiere, in der Lage zu sein scheinen, Individuen zu unterscheiden, ist entsprechend kein Argument für deren Fähigkeit, mit sortalen Gegenständen umzugehen: Diese Wesen reagieren auf hinreichend genaue Attrappen ebenso. Dass es unter natürlichen Bedingungen in der Regel nur ein Individuum (etwa das versorgende Elternteil) gibt, das für die entsprechenden Wahrnehmungen sorgt, macht diesen Verhaltenstypus zwar ebenso regelmäßig erfolgreich, hat aber mit dem anspruchsvollen Begriff des individuierten Gegenstands nur sehr wenig zu tun.

lich oder zeitlich auszuweiten, sondern darum, die Aufmerksamkeit zugleich auf eine (oder mehrere) ganz andere, mit der aktuellen Situation unverbundene Raumzeit-Blase(n) richten zu können.

Wie könnten wir im oben erwähnten Beispiel feststellen, ob die Schimpanse Moja zu einer solchen Distanzierungsleistung in der Lage ist? Nur so ließe sich ja die Vermutung des Assistenten bestätigen, sie habe ihre Aufmerksamkeit sowohl mit den Linien auf dem Blatt Papier als auch mit dem ASL-Symbol für ‚Vogel‘ demselben nicht-anwesenden Individuum zugewandt? Nun, wie würden wir das bei einem unserer Artgenossen entscheiden? Sicherlich genügt es nicht, etwa auf unsere Beobachtung eines geistesabwesenden Gesichtsausdrucks zu bestimmten Zeiten hinzuweisen. Auch das direkte Umgehen mit einem Gegenstand kann keine unmittelbar schlüssigen Beweise liefern, dass dieser als ein sortaler Gegenstand erlebt wird, bleibt jedes solche Verhalten doch an den je aktuellen Kontext gebunden. Vielmehr würden wir den Betreffenden fragen und erwarten, dass er sich uns gegenüber aktiv *darstellen* würde als jemand, der seine Aufmerksamkeit auf einen anderen Kontext richten kann.

Faktisch bilden also allein bestimmte komplexe Zeichenhandlungen das Kriterium, genauer: solche Zeichenhandlungen, die einen propositionalen Gehalt aufweisen. Nach allem, was wir wissen, lässt sich prinzipiell nur kommunikativ bestimmen, ob jemand mit sortalen Gegenständen umgehen kann, da es nur so, nämlich gemeinsam und durch wechselseitige Kontrolle, möglich ist, stabilen Zugang zu einem nicht-anwesenden Kontext zu erreichen. Der Verweis auf einen nicht-anwesenden Kontext erfolgt, indem ein Kommunikationspartner sich dem andern gegenüber darstellt als jemand, der seine Aufmerksamkeit auf jenen Kontext (und nicht nur den tatsächlich anwesenden) gerichtet hat. Sich darstellen als jemand, der seine Aufmerksamkeit auf einen bestimmten Kontext richtet, ist daher eine kommunikative Handlung, die zentral ist für das Umgehen mit sortalen Gegenständen wie für das Bildvermögen; eine Handlung, die im Weiteren ‚Kontextbildung‘ genannt sei.<sup>16</sup>

## 7. Ähnlichkeit, handlungstheoretisch gefasst

Tatsächlich hängt auch die Fähigkeit, Ähnlichkeit erkennen zu können, ganz wesentlich an dem mit der Kontextbildung ermöglichten Vergleich zwischen ganz verschiedenen Kontexten, womit sich ein Weg eröffnet, die eigentümliche Kombination von Zeichenverwendung und Ähnlichkeit, die für den Bild-

<sup>16</sup> Der Ausdruck wurde in Anlehnung an die „*space builder*“ des Linguisten Fauconnier (Gilles FAUCCONNIER: *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge 1985) gebildet.

gebrauch so typisch zu sein scheint, genauer unter die Lupe zu nehmen. Es wäre sehr naiv zu meinen, Ähnlichkeit sei einfach eine objektive Relation zwischen vorgegebenen Gegenständen, eine Art abgeschwächte Gleichheit oder Isomorphie. Denn was überhaupt als Gegenstände vorkommt, wird ja nur mit Bezug auf ein Wesen, das mit ihnen umgeht, sie wahrnimmt und auf sie reagiert, bestimmt. Für einen handlungstheoretisch gefassten Begriff der Ähnlichkeit ist es unumgänglich, dem Hier und Jetzt eine alternative Verhaltenssituation gegenüberstellen zu können, da von Ähnlichkeit im eigentlichen Sinn nur gesprochen werden kann, wo es um sortale Gegenstände geht und der erst mit ihnen gegebene Unterschied zwischen aktuellem Schein und darüber hinaus gehendem Sein zur Verfügung steht.

Wird aber mit einer solch strengen Verbindung nicht zuviel behauptet? Reagieren nicht schon selbst recht einfache Tiere auf Ähnlichkeit – etwa im Zusammenhang mit Mimikry oder der Verwendung von Attrappen – obwohl hier sicherlich noch keine Befähigung, mit sortalen Gegenständen umzugehen, besteht? So legt die durch Plinius d. J. erzählte Anekdote von dem antiken Maler Zeuxis nahe anzunehmen, dass Vögel, denen wir gemeinhin nicht zuschreiben, individuierte Gegenstände zu beherrschen, doch Ähnlichkeit erfassen (vgl. Abb. 3). Plinius zufolge hatte Zeuxis seine Darstellung von Trauben so ähnlich gestaltet, dass einige Vögel herbei flogen, um nach den scheinbaren Früchten zu picken. Entscheidend für die handlungstheoretische Bestimmung des Begriffs der Ähnlichkeit ist, dass das Verhalten der Vögel als ein der aktuellen Verhaltenssituation nicht adäquates Verhalten betrachtet wird. Das Verhalten würde jedoch, so denken wir Beobachter, zu einem ganz anderen Kontext durchaus gut passen – einer Verhaltenssituation, in der an Stelle des Bildträgers tatsächlich Futter wäre. Das Verhalten der Vögel zeigt uns, mit anderen Worten, dass sie sich vom falschen Schein täuschen lassen. Dass er falsch ist, ist *uns* klar. Die Vögel durchschauen hingegen ihr Verhalten nicht selbst als Täuschung. Zwar werden sie sich auch nach einiger Zeit ‚enttäuscht‘ abwenden. Doch bleibt für sie dann an der Stelle, an der sie eben noch Trauben wahrgenommen hatten (*B*-Wahrnehmung in Abb. 3), nicht etwas zurück, was an Trauben erinnert, sondern lediglich etwas, was mit Trauben gar nichts mehr zu tun hat (*A*-Wahrnehmung). Dass eine gewisse Wahrnehmung bei einem anderen Wesen vorliegt, erkennen wir nur an dessen zugehörigen Reaktionen. Auch dass Gegenstände in einigen Aspekten gleich, in anderen ungleich sind, ist daher nur relativ zu den entsprechend gleichen oder verschiedenen Reaktionen des betrachteten Wesens bestimmt.<sup>17</sup> Ferner ist es charakteristisch

<sup>17</sup> Dieser Aspekt wurde in Abbildung 3 natürlich nicht adäquat dargestellt: Es wäre recht schwierig, die entsprechenden Verhaltenskomplexe in einer einfachen Skizze deutlich zu markieren. Die aus dem Comic wohl bekannten ‚Denkblasen‘ vertreten die Reaktionen; doch sollte klar sein, dass sie lediglich eine – leicht falsch zu verste-

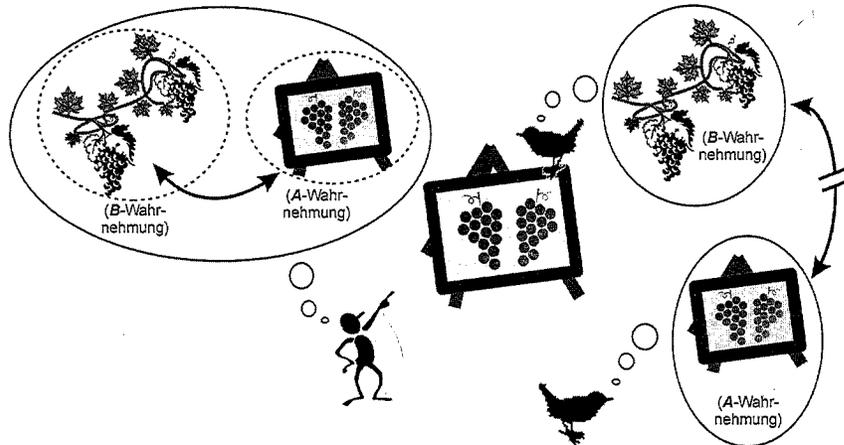


Abb. 3: Um eine Täuschung (Ähnlichkeit) als solche zu erkennen, muss man vergleichen können

für das Wahrnehmen auf dieser Komplexitätsstufe, dass es lediglich in der einen oder anderen Hinsicht Gleichartiges – insofern nämlich dasselbe Verhaltensmuster damit verbunden ist – von anderem unterscheidet, nicht aber Individuelles differenziert. Wahrnehmung heißt hier stets, Gleichartiges (also Ähnliches) zusammen zu gruppieren. Daher kann Ähnlichkeit selbst noch nicht erkannt werden.

Anders als Vögel verfügen hingegen Menschen mit den sortalen Begriffen prinzipiell über die Möglichkeit, in einer *trompe l'œil*-Situation nicht einfach nur entweder fehlerhaft auf die Täuschung hereinzufallen oder gar keinen Zusammenhang zwischen Bildträger und Abgebildetem herzustellen. Vielmehr gelingt es ihnen, sich zugleich anwesende Darstellung (A-Wahrnehmung in Abb. 3, auf aktuellen Kontext bezogen) und abwesendes Dargestelltes (B-Wahrnehmung, auf einen anderen Kontext bezogen) zu vergegenwärtigen und beides als zwei verschiedene, aber aufeinander verweisende Entitäten zu begreifen. Im Falle der Ähnlichkeit zwischen einem Ausschnitt des Bildträgers und einem damit abgebildeten Objekt beruht das Zuschreiben der Ähnlichkeit darauf, dass zwischen dem aktuellen Kontext mit dem Bildträger und dem nur fiktiven Kontext mit dem Abgebildeten eine Beziehung hergestellt wird, also eine Relation zwischen zwei Verhaltenssituationen. Für das Beispiel heißt das, dass es auch für uns gerade an den spontanen ‚falschen‘ Reaktionen, die das Traubenwahrnehmen charakterisieren, liegt, dass die Traubendarstellung nicht einfach als gegenwärtige Beeren wahrgenommen, sondern als

hende – Verkürzung des eigentlichen begrifflichen Zusammenhangs zwischen (zugeschriebener innerer) Wahrnehmung und (beobachtbarem) Verhalten sind.

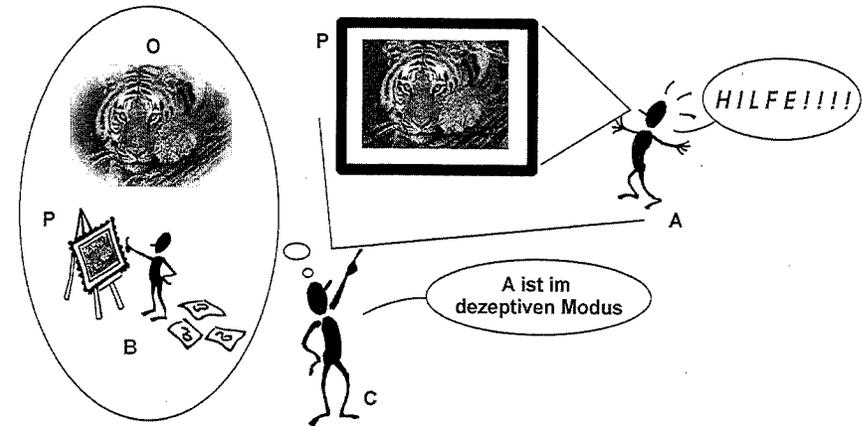


Abb. 4: Der dezeptive Verhaltensmodus

den Früchten lediglich ähnlich erkannt wird: Nur insofern einerseits unwillkürlich eine Traubenwahrnehmung (mit den damit spontan verbundenen Reaktionen) ausgelöst wird – wie schwach auch immer – kann davon die Rede sein, dass einem Wesen irgendein wahrgenommener Gegenstand als solchen Trauben ähnlich erscheint. Nur insofern andererseits der Bildträger als etwas wahrgenommen wird, was jedenfalls keine Trauben sind, kann die spontane Täuschung darüber hinaus durchschaut werden.

Im eigentlichen Sinne soll also nur dann von Ähnlichkeit die Rede sein, wenn ein Wesen (a) im aktuellen Kontext auf einen Gegenstand spontan mit einem Verhalten zu reagieren geneigt ist, das nicht zu diesem Gegenstand (und damit dem aktuellen Kontext) passt, wohl aber zu einem Kontext mit einem anderen Objekt, und (b) erkennt, dass es sich um eine solche Verwechslung handelt, es also eine Verbindung zwischen den beiden beteiligten Kontexten herzustellen in der Lage ist.<sup>18</sup>

Damit ergibt sich auch eine genauere Bestimmung des Begriffs der wahrnehmungsnahen Zeichen (und damit der Bilder) aus handlungstheoretischer Perspektive: Betrachten wir dazu zunächst die grundlegendere Situation der Täuschung, wie sie schon bei Zeuxis' Vögeln auftrat: Dabei wird der (potentielle) Bildträger mit dem Abgebildeten spontan verwechselt, d.h. entspre-

18 Da es aber umgangssprachlich zu nahe liegt und daher kaum zu vermeiden ist, dass auch bei den Vögeln des Zeuxis und ähnlichen Fällen von ‚Ähnlichkeit‘ gesprochen wird, ist es zumindest sinnvoll, die beiden Fälle durch Indizes voneinander zu unterscheiden. In Jörg R.J. SCHIRRA: *Foundation of Computational Visualistics*. Wiesbaden 2005, werden beispielsweise entsprechend ‚Ähnlichkeit<sub>a</sub>‘ beim Zuschreibung einer Täuschung bei Lebewesen, welche die Täuschung selbst nicht durchschauen können, und ‚Ähnlichkeit<sub>b</sub>‘ im anderen Fall verwendet. Anders gewendet wird der Begriff Ähnlichkeit<sub>a</sub> beim Übergang zu sortalen Gegenständen differenziert in Ähnlichkeit<sub>a</sub> einerseits und Identität andererseits.

chende unmittelbare Verhaltensweisen werden ausgelöst, ohne dass der Fehler bemerkt wird (vgl. Abb. 4). Wir sprechen vom *dezeptiven* Modus. Natürlich kann nur jemand einem von ihm beobachteten Wesen den dezeptiven Verhaltensmodus zuschreiben, der selbst in der Lage ist, die Täuschung zu durchschauen.

Erkennt jemand andererseits einen (potentiellen) Bildträger als wesentlichen Teil einer kommunikativen Situation, so handelt es sich um den *symbolischen* Modus. Derjenige begreift also, dass da ein Gegenstand anwesend ist, mit dessen Hilfe ein Sender intendiert, die Aufmerksamkeit eines Empfängers auf etwas zu richten. Bei den uns interessierenden komplexeren Formen handelt es sich dabei allerdings um etwas nicht zugleich Anwesendes, das mittels des Zeichenträgers *vergegenwärtigt* werden soll. Der symbolische Modus zeichnet sich insbesondere dadurch aus, dass man auf das Repräsentierte, etwa einen Tiger, da es ja nur symbolisch gegenwärtig ist, nicht so zu reagieren braucht wie auf seine tatsächliche Anwesenheit (vgl. Abb. 5). Da Ähnlichkeit für den symbolischen Modus keine Rolle zu spielen braucht, gilt er gleichermaßen für alle Zeichenarten, nicht nur für Bilder.

Beim Gebrauch von wahrnehmungsnahen Zeichen soll nun eine Täuschung auf systematische Weise mit einem Zeichengebrauch zusammenspielen: Der Zeichenträger soll dem Repräsentierten als ähnlich erkannt werden. Er sollte also eine mehr oder weniger starke spontane Verwechslungsreaktion auslösen können, was dem dezeptiven Modus entspricht. Die inadäquate Reaktion tritt allerdings – wegen der Einbettung in den symbolischen Modus – in der Regel nicht nach außen. Sie wird nur wirksam in dem, was mit der Zeichenhandlung (vermutlich) intendiert wird (vgl. Abb. 6). Das ist der *immersiven* Modus, bei dem man die Täuschung sowohl erlebt als auch durchschaut – d.h. als Ähnlichkeit begreift – und so als Grundlage für eine Zeichenverwendung nehmen kann. Darüber hinaus erfüllen die so bestimmten Zeichen das Kriterium der Kontextbildung: Insofern nämlich die Zeichenhandlung so verstanden wird, dass mit ihr auf die Situation aufmerksam gemacht werden soll, in der die in der Täuschung inadäquate Reaktion adäquat wäre (im Falle eines Bildes also auf eine Situation, in der das Abgebildete anwesend wäre). Es liegt daher nahe, die Kontextbildung generell als die grundlegende kommunikative Funktion des Bildgebrauchs zu betrachten.

Es gibt noch einen vierten Modus, auf den auch Jonas im Zusammenhang mit den Graden der Ähnlichkeit hinweist: Er deckt insbesondere den oben erwähnten Komplex „Emanzipation von der Wörtlichkeit“ (Punkt 3c in Abschn. 4) ab: Neben seinem direkten („wörtlichen“) Gebrauch kann ein Zeichen nämlich prinzipiell auch exemplarisch benutzt werden. Es wird dann dazu verwendet, den Kommunikationspartner (oder sich selbst in der Rolle eines anderen) am Beispiel auf Aspekte der entsprechenden Zeichenverwendung aufmerksam

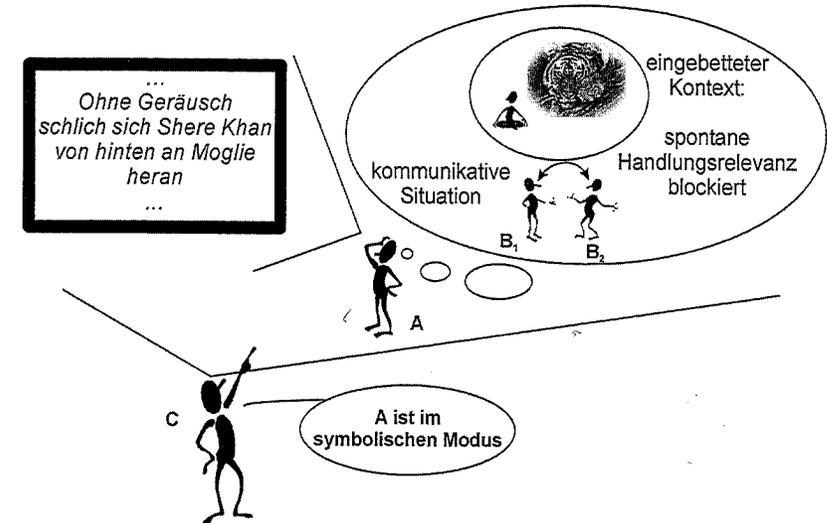


Abb. 5: Der symbolische Verhaltensmodus

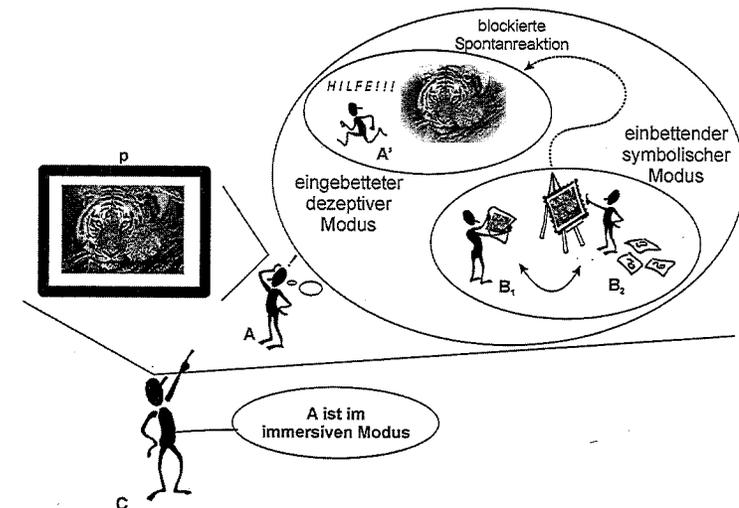


Abb. 6: Der immersive Verhaltensmodus

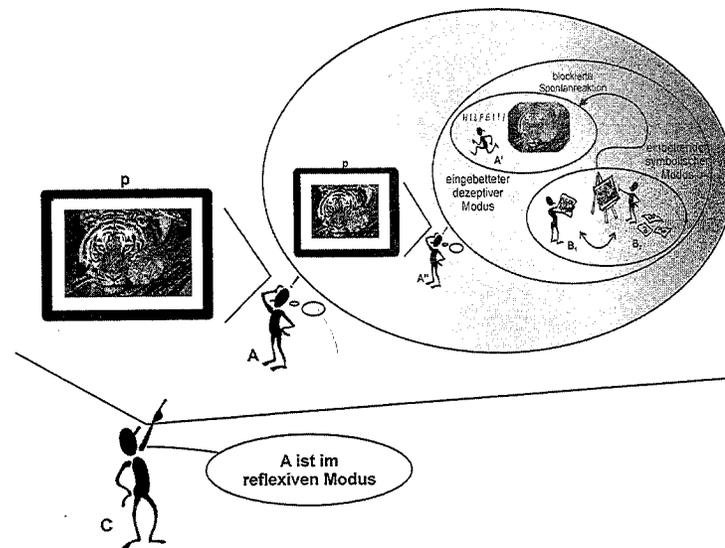


Abb. 7: Der reflexive Verhaltensmodus  
(Der ‚Nebel‘ soll andeuten, dass hier die Aufmerksamkeit exemplarisch auf die syntaktischen Aspekte gelenkt wird)

zu machen. Auf diese Weise kann insbesondere auf defiziente Gebrauchsweisen eingegangen werden. Dieser *reflexive* Modus soll uns hier nicht weiter interessieren (vgl. Abb. 7). Auf Bilder bezogen spielt er vor allem eine Rolle bei Bildzitationen, wie sie z.B. in bildwissenschaftlichen Diskursen aufzutreten pflegen. Das trifft etwa auf Abbildung 1 (Moja's ‚Bild‘) zu. In der bildenden Kunst tritt dieser Modus ebenfalls in Erscheinung. Auch für den reflexiven Gebrauch eines Bildes spielt die Fähigkeit zur Kontextbildung eine zentrale Rolle, wird doch in ihm die Aufmerksamkeit gewissermaßen von außen auf die aktuelle Situation des Bildverwendens gerichtet.

Insgesamt wurde damit ein im Sinne des *linguistic turn* ganz auf Verhalten aufbauender Vorschlag für den Bildbegriff skizziert. Bestimmt über das Einbetten eines spontan ausgelösten inadäquaten Verhaltens in eine Zeichenhandlung, die auf die Situation verweist, in der das Verhalten adäquat wäre, steht dieser Vorschlag der populären Fassung gegenüber, bei der letztlich ein doch recht unklarer Begriff der Ähnlichkeit einfach vorausgesetzt und die enge begriffliche Verflechtung von Ähnlichkeit, Identität und sortalen Gegenständen ignoriert wird.

## 8. Ansatz einer begriffsgenetischen Begründung des Bildbegriffs

Die handlungstheoretische Bestimmung des Bildbegriffs führt ganz zwanglos auch zu dem wesentlich weiter gefassten Ansatz seiner begriffsgenetischen Begründung, der hier – gleichsam als Forschungsprogramm – nur kurz umrissen sei. Einerseits muss sich der Bildanthropologe einlassen auf verschiedene Stufen der Komplexität der Begriffe von Wesen, die Zeichen gebrauchen können. Andererseits soll Wahrnehmungsnähe als Ziel einer spezifischen Differenz verwendet werden: Bildanthropologen müssen sich demnach auch verschieden komplexen Begriffsfeldern um die Begriffe ‚Aktivitätsträger, die in einem mehr oder weniger anspruchsvollen Sinn wahrnehmen können‘ zuwenden. Ausgangspunkt einer solchen begriffsgenetischen Untersuchung sollten jeweils die Begriffsfelder für Wesen sein, bei denen noch nicht von Wahrnehmung bzw. von Zeichengebrauch im elementarsten Sinn gesprochen werden kann. Derartige Stufentheorien sind in der Ethologie und der Sprachphilosophie vorzufinden<sup>19</sup> und wurden bereits ansatzweise in den beiden vorigen Kapiteln ins Spiel gebracht.

Ziel der begriffsgenetischen Betrachtungen soll es sein, aus den beiden Abfolgen – der semiotischen und der wahrnehmungstheoretischen – die (minimale) Stufe zu bestimmen, bei der die eigentümliche Kombination von Zeichengebrauch und Wahrnehmungsnähe auftritt, die uns zumindest für darstellende Bilder als charakteristisch gilt (Abb. 8). Das damit gegebene minimale Feld mit einem Begriff für Bilder benutzende Wesen versammelt genau die wesentlichen Bestimmungsstücke des immersiven Modus und nichts anderes. Von dem minimalen Feld ausgehend könnten komplexere Felder, die höhere Stufen des Bildvermögens charakterisieren, abgeleitet wie auch metaphorische Assoziationen rekonstruiert werden.<sup>20</sup> Ferner sollte damit die Beziehung zu jener Stufe geklärt werden können, auf welcher der Gebrauch propositionaler Sprache möglich wird – sie muss ohnehin Teil der ‚semiotischen‘ Stufenfolge sein. So wäre klar, ob das Sprachvermögen eine unumgängliche Voraussetzung des Bildvermögens ist oder umgekehrt; oder ob beide sich wechselseitig bedingen, so dass von Wesen, die nur über eines der beiden

19 Vgl. etwa Arno Ros: *Materie und Geist. Eine philosophische Untersuchung*. Paderborn 2005.

20 Es ist natürlich kein Zufall, dass sich die Bildwissenschaft historisch gespalten präsentiert: in einen semiotischen Zweig, dessen Proponenten den Bildbegriff vor allem im Rahmen der Zeichentheorie analysierten (Nelson GOODMAN: *Languages of Art. An Approach to a Theory of Symbols*. Indianapolis 1968), und einen häufig ‚phänomenologisch‘ benannten Zweig, dessen Befürworter sich vor allem für die Besonderheiten des Wahrnehmens beim Bildgebrauch interessierten (Ernst H. GOMBRICH: *Art and Illusion*. Princeton 1960); vgl. auch Dominic LOPES: *Understanding Pictures*. Oxford 1996. Die begriffsgenetische Synthese verspricht diesen allzu häufig vehement geführten Richtungsstreit zu überwinden.

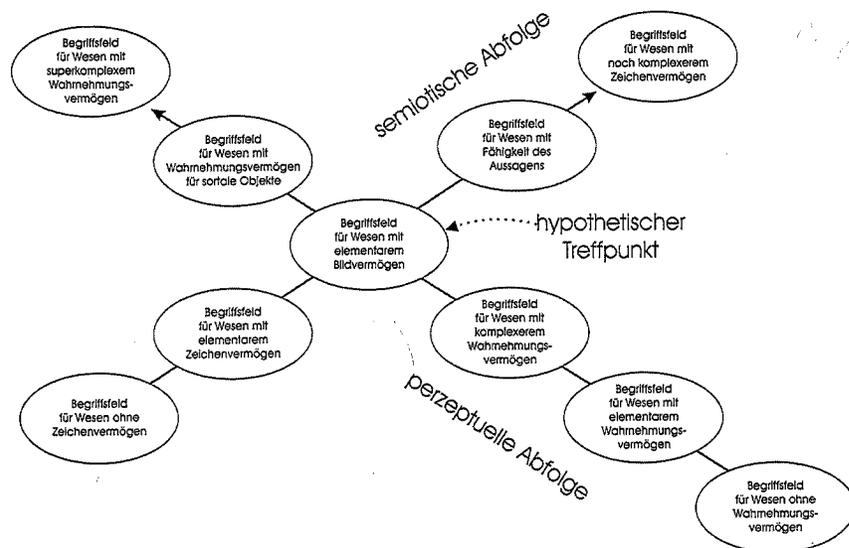


Abb. 8: Hypothetisches Schema der beiden Abfolgen von Begriffsfeldern

Die geraden Pfeile deuten die begriffsgenetische Konstitutionsrelation an. An dem – hier rein spekulativ eingezeichneten – Treffpunkt können die Wahrnehmungsfähigkeiten mit dem Zeichenvermögen zum immersiven Modus kombiniert werden. Im dargestellten Fall wäre das Zeichenvermögen konstitutiv für die Fähigkeit, Aussagen zu verwenden.

Vermögen verfügen, prinzipiell nicht die (vernünftige) Rede sein kann. Man darf Jonas in seinem *homo pictor*-Aufsatz wohl so verstehen, dass er dazu tendiert, der Bildverwendung das Primat vor dem Sprachvermögen zuzuordnen – eine Möglichkeit, die auch in Abbildung 8 skizziert erscheint.<sup>21</sup> Eine solche Relation wirkt auf den ersten Blick auch durchaus plausibel, gilt propositionale Sprache doch zu Recht als das mächtigere und komplexere Kommunikationswerkzeug.

Ein Gegenargument ergibt sich allerdings aus der handlungstheoretischen Ableitung des Bildbegriffs: dass nämlich einerseits Ähnlichkeit nur erkennen kann, wer mit sortalen Gegenständen umzugehen in der Lage ist, und andererseits das Umgehen mit sortalen Gegenständen von der Fähigkeit, beliebige Kontexte evozieren zu können, abhängt – einer Fähigkeit, die auf eindeutige Weise nur mittels propositionaler Sprache nachgewiesen werden kann. Da-

21 Am deutlichsten wohl in folgender Passage: „Die *adequatio imaginis ad rem*, die der *adequatio intellectus ad rem* vorangeht, ist die erste Form theoretischer Wahrheit“ (Hans JONAS: Die Freiheit des Bildens. *Homo pictor* und die differentia des Menschen. In: Zeitschrift für Philosophische Forschung 15 (1961), S. 161–176, hier S. 40. Zitiert nach dem Wiederabdruck in DERS.: Zwischen Nichts und Ewigkeit. Zur Lehre vom Menschen. Göttingen 1987, S. 26–43).

mit wäre also umgekehrt das Sprachvermögen eine Voraussetzung für Bildvermögen.<sup>22</sup> Solche wechselseitigen Abhängigkeiten sprechen sehr dafür, dass es sich tatsächlich um Fähigkeiten handelt, die in ganz engem begrifflichen Zusammenhang stehen und nicht unabhängig voneinander gedacht werden sollten.

Natürlich sprengte eine derartige präzise ausgeführte begriffsgenetische Untersuchung bei weitem den Umfang unserer Abhandlung. Daher soll abschließend nur noch der Aspekt der wechselseitigen Abhängigkeit von Bild- und Sprachvermögen etwas näher beleuchtet werden. Zuvor lohnt es, sich kurz die Besonderheiten des menschlichen Sprachvermögens zu vergegenwärtigen. Wie bereits erwähnt stehen hier Aussagesätze im Fokus des Interesses, oder allgemeiner: solche Zeichenhandlungen, die neben der illokutionären Funktion einen propositionalen Gehalt aufweisen. Diese Aufteilung trägt der Tatsache Rechnung, dass man, indem man Sprache benutzt, in der Regel auch noch etwas anderes tut als eben nur zu sprechen.<sup>23</sup> Man warnt oder verspricht, fragt oder befiehlt, behauptet oder bezweifelt, um nur einige Beispiele für illokutionäre Funktionen zu nennen. Von diesem Aspekt, der die mit der Sprachhandlung vollzogene Interaktion charakterisiert, ist der Sachbezug zu unterscheiden: das, wovon man warnt, was man verspricht, erfragt oder befiehlt. Der den Sachbezug vermittelnde propositionalen Gehalt wird als aus zwei Arten von Teilhandlungen zusammengesetzt verstanden: 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; mit der (oder den) *Nomination(en)* versucht er verständlich zu machen, auf welchen einzelnen Gegenstand (oder Gegenstände) er sich dabei beziehen möchte: Diese Gegenstände müssen dabei beiden Kommunikationspartnern bereits bekannt – in ihrem gemeinsamen ‚Diskursuniversum‘ vorhanden – sein. Das jeweilige Diskursuniversum ist aber nun nichts anderes als ein bestimmter von allen Gesprächspartnern geteilter Kontext.

Aussagen sind daher *kontextrelativ*: Wenn das zugehörige Diskursuniversum unbekannt ist, bleibt ein Aussagesatz wesentlich unverständlich. Insbesondere die Nomination kann nur erfolgreich durchgeführt werden, wenn klar ist, welche Menge von Gegenständen zur Auswahl steht. Andererseits aber sind Aussagen *situationsunabhängig*: Man kann von jedem beliebigen situativen Kontext aus Aussagen vollziehen, die sich auf irgendeinen anderen Kon-

22 Das folgt jedenfalls, wenn Bildverwendungen vom rein dezepativen Modus unterschieden werden sollen.

23 John L. AUSTIN: Zur Theorie der Sprechakte. *How to do things with Words*. Stuttgart 1972.

text beziehen. Beide Charakterisierungen der Aussage bedingen sich wechselseitig, denn nur, indem explizit auf einen Kontext verwiesen wird, kann vom *aktuellen* Kontext unabhängig geredet werden. Aussagen bedürfen, mit anderen Worten, notwendigerweise eines Aktes der Kontextbildung. Zum Beispiel dienen sprachliche Orts- oder Zeitangaben der Kontextbildung. Aber auch hypothetische oder fiktive Kontexte können verbal gebildet werden, etwa durch Verweise auf Texte literarischer Fiktion: „In Johnsons Roman ‚Jahrestage‘ ist Lisbeth Mutter von Gesine.“ Wen der Sprecher mit „Gesine“ oder „Lisbeth“ meint, ist nur relativ zu dem angegebenen literarisch-fiktiven Kontext verständlich, die Aussage daher nur im Bezug auf ihn verifizierbar.<sup>24</sup>

Während Wesen mit nur einfacher Zeichensprachkompetenz sich immer auf Aspekte der Äußerungssituation beziehen,<sup>25</sup> zeigt sich eine Differenzierung in Prädikation und Nomination, wie sie Aussagen charakterisiert, überhaupt nur dann nützlich, wenn man auf Sachverhalte zu sprechen kommen möchte, die nicht ohnehin in der aktuellen Verhaltenssituation (oder als Ziel aktueller Absichten) gegenwärtig sind. Sehen wir im folgenden von Aussagen über den unmittelbaren Äußerungskontext ab und beschränken uns zunächst auch darauf, rein sprachliche Kontextbildungen zu betrachten: Um das in der Aussage behauptete Zutreffen bestimmter Begriffe auf die angegebenen Einzelgegenstände empirisch einlösen zu können, müsste man die aktuelle Verhaltenssituation verlassen und sich in den spezifizierten Kontext begeben. Will oder kann ich den erwähnten Kontext nicht zum aktuellen Kontext machen (etwa bei fiktiven Kontexten), bleibt nur übrig, durch logische Schlüsse weitere Aussagen abzuleiten und auf Konsistenz mit dem bereits über jenen Kontext Gewussten zu untersuchen. Empirisch kann das Gelten der Behauptung so jedoch nicht nachgeprüft werden.

Wird aber die Kontextbildung durch das Präsentieren eines Bildes ausgeführt, so wird eine zusätzliche Verhaltenssituation heraufbeschworen, die – im dezeptiven Modus – wahrgenommen, d.h. auf die spontan reagiert werden kann: Die entsprechenden senso-motorischen Testroutinen der angesprochenen Unterscheidungsgewohnheiten sind (zumindest partiell) direkt anwendbar. Dieser Kontext ist nicht, wie die sprachlich gebildeten Kontexte, völlig abgetrennt von der aktuellen Situation des Zeichengebrauchs. Beide erscheinen

24 Die aktuelle Verhaltenssituation spielt sicher eine ausgezeichnete Rolle, denn nur bei den Aussagen, die sich darauf beziehen, greift die referentielle Verankerung von Nomination und Prädikation unmittelbar und die senso-motorischen Komponenten der verwendeten Begriffe – also die zugehörigen Unterscheidungsgewohnheiten – können ausgespielt werden. Kontextbildungen, die wie das Satzadverbial „in Prag“ auf Orte verweisen, geben implizit eine Methode an, wie der gemeinte Kontext in die aktuelle Verhaltenssituation überführt werden könnte, um dann die referentielle Verankerung des propositionalen Gehalts der Äußerung durchzuführen.

25 Vgl. etwa Ernst TUGENDHAT: Vorlesungen zur Einführung in die sprachanalytische Philosophie. Frankfurt am Main 1976, S. 208ff.

vielmehr teilweise miteinander ‚verschmolzen‘. Während also verbale Kontextbildung Gegenstände, auf die durch Nominationen verwiesen wurde, nur *logisch vergegenwärtigt*, kann piktoriale Kontextbildung sie *empirisch vergegenwärtigen*.

Stellen wir uns in einem Gedankenexperiment Wesen vor, die zwar Aussagen machen können, aber prinzipiell keine wahrnehmungsnahen Zeichen kennen. Auch zu sortalen Objekten hätten jene Wesen lediglich sprachlich Zugang: Die Wahrnehmung bietet ihnen nur jeweils die Erscheinungen in *einem*, dem aktuellen Kontext, die sie bloß durch Sprache mit Erscheinungen in anderen Verhaltenssituationen in Beziehung zu setzen imstande wären, also auf eine Art, die empirische Überprüfung nicht zulässt.

Auf welche Weise sind diese hypothetischen Wesen wohl dazu gekommen, sich überhaupt auf nicht-aktuelle Kontexte zu beziehen? Welcher Weg führte sie zur Objektkonstitution – d.h. zum Gebrauch von Gegenständen mit sortaler Individuation? Beziehungsweise: mit welcher Überlegung können wir auf rationale Weise zum Begriff eines solchen Wesens kommen? Denn zum Übergang vom Begriffsfeld der Wesen, die sich lediglich über Sachverhalte in der je aktuellen Situation – empirisch gegenwärtig – verständigen können, zu dem Begriffsfeld der Wesen im Gedankenexperiment – mit einer Kontextbildung, die rein logisch vergegenwärtigen kann – klafft uns eine ausgesprochen breite Lücke.

## 9. Die initiale Kontextbildung und der Begriff des inneren Bildes

Wäre es denkbar, dass das Vermögen, situationsunabhängig – mit Aussagen – zu kommunizieren, von den durch Bilder vermittelten Kontexten nicht nur ‚unter anderem‘ Gebrauch macht, sondern überhaupt erst durch das täuschende Potential bildartiger Vorläufer entstehen konnte? Die Frage zielt ab auf die *initialen Kontextbildung*, dem ursprünglichen das ‚Hier und Jetzt‘ transzendierenden Akt. Zwar bleibt die Verwendung eines wahrnehmungsnahen Zeichens mehrdeutig. Man kann ihm ja auch bloß im dezeptiven Modus gegenübertreten und nicht merken, dass es um einen *anderen* Kontext gehen soll. Aber dieser Mangel mag sich bei der Einführung der Kontextbildung als ein Vorteil entpuppen.

Gehen wir aus von Wesen, die nur auf ihren je aktuellen Verhaltenskontext *a* Zugriff haben und daher einem (potentiellen) Bildträger *B* höchstens im dezeptiven Modus – Verwechslung von Situation *a* mit einer Situation *b* – gegenübertreten können. Dieses Verhalten – etwa ein Balzverhalten angesichts einer Attrappe – ist noch nicht ohne weiteres kommunikativ zu gebrauchen. Der symbolische Modus läge erst vor, wenn das Wesen (*i*) in der Lage wäre, ande-

ren jenes Verhalten auch *vorzuführen*: also einem anderen „körperlich zu zeigen, dass es dabei ist, etwas zu verwechseln (ohne das selbst zu merken)“<sup>26</sup>; wenn es (ii) auf dieses eigene Vorführverhalten auch selbst so reagieren würde, wie ein Empfänger, da seine Bedeutung ansonsten für Sender und Empfänger verschieden bleibt; und wenn (iii) das Vorführen von Verhaltensweisen, wie es für den Vollzug einfacher zeichensprachlicher Handlungen wesentlich ist, samt der dazugehörigen Antwortreaktionen *verinnerlicht* werden könnte: Dann genügt eine sehr schwache Aktivierung der entsprechenden Nervenbahnen, eine fast unmerkliche Änderung des zugehörigen Muskeltonus, die zwar kaum zu einer von außen noch unterscheidbaren Aktivität führt, doch über das propriozeptorisch vermittelte Körperbewusstsein innerlich wirksam bleibt.<sup>27</sup>

Die Besonderheit des kommunikativ eingesetzten dezeptiven Verhaltens ist nun, dass die Antwortreaktion zweigeteilt ausfallen kann, je nachdem, ob der Empfänger ebenfalls der Täuschung erliegt oder nicht. Praktisch kann man sich durchaus vorstellen, dass etwa bestimmte natürliche Felsformationen immer wieder zu Verwechslungen mit einem Fressfeind Anlass geben. Signal-sprachlich miteinander kommunizierende Gruppen könnten sich dort daran gewöhnen, entsprechende Warnsignale eines Artgenossen in diesem Kontext zu ignorieren oder gar prompt Entwarnsignale zu äußern. Damit wäre ein Komplex aus einem Vorführverhalten (im dezeptiven Modus) und *zweierlei* Reaktionen darauf gegeben: Der Sender reagiert mit seinem Signal auf den für ihn aktuellen Kontext *b*, auf den auch das Verhalten eines Empfängers, der auf das Warnsignal normal reagiert, gerichtet ist, während schließlich für einen Empfänger, der das Signal auf die erlernte Weise ignoriert, Kontext *a* der aktuelle ist (Abb. 9). Das Verinnerlichen des ursprünglich externen Zeichenverhaltens in einer solchen potentiellen Täuschungssituation kann dann entsprechend *beide* Reaktionsoptionen bei dem reflexiven Sender innerlich auslösen und so das Fundament für eine Relation zwischen den beiden beteiligten Kontexten für dieses Wesen bilden. Dann können wir auch davon sprechen, dass

26 Als konkretes Beispiel aus der ethologischen Forschung mag Folgendes dienen: Manche Affen verwenden akustische Warnsignale, wenn bestimmte Feinde wahrgenommen, also als aktuell anwesend erkannt werden, etwa große Greifvögel oder Würmeschlangen. Nun ist es plausibel, dass ein Individuum einer dieser Affenarten das Warnsignal ausstößt, wenn es auf eine (etwa von einem Ethologen entsprechend platzierte) Schlangentrappe aus Gummi stößt und so seine Gruppe zur Flucht veranlasst, obwohl tatsächlich keine Schlange anwesend ist.

27 Präziser gesagt geht es darum, dass diese Verhalten „sich hier aus den kommunikativen Zusammenhängen, in die sie auf der Stufe des Vollzugs zeichensprachlicher Handlungen normalerweise eingebettet sind, herauslösen, und zur Fähigkeit einer allein für sich zu vollziehenden leiblichen Selbstvergegenwärtigung entwickeln“ können (ROS: Materie (wie Anm. 19), S. 591). Siehe insbesondere auch George Herbert MEAD: *Mind, Self, and Society*. Chicago 1934 und Ernst TUGENDHAT: *Vorlesungen zur Einführung in die sprachanalytische Philosophie*. Frankfurt am Main 1976, V. 13.

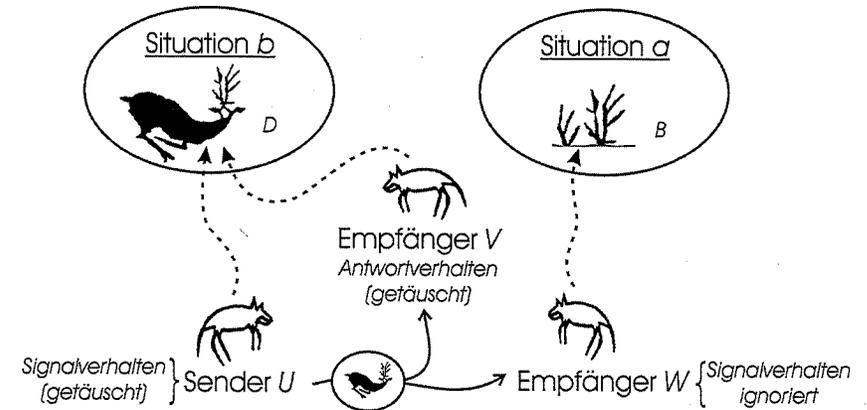


Abb. 9: Schema zur Signalkommunikation in einer Verwechslungssituation (Hirsch / Geäst)

sich dieses Wesen anderen und sich selbst gegenüber darstellt als ein Wesen, das etwas wahrnimmt, was so gar nicht anwesend ist. Dabei bleibt dieses Verhalten noch ganz abhängig von der wirklichen Anwesenheit des Gegenstands *B* im aktuellen Kontext, einem Gegenstand, der also tatsächlich gerade wahrgenommen werden kann und dabei leicht mit etwas anderem (*D*) zu verwechseln (d.h. ihm ähnlich) ist.

Die Überlegung führt zum Begriff eines Wesens, dem eine Vorform der piktorialen Kontextbildung gelingt: Es hat den Bildträger im immersiven Modus gebraucht. Die damit knapp skizzierte Ableitung der initialen Kontextbildung im Gefolge der begriffsgenetischen Kopplung von dezeptivem und symbolischem Modus zum immersiven Modus hinsichtlich eines potentiellen Bildträgers *B* ist allerdings noch nicht stabil, denn kein Mechanismus verhindert den Rückfall in den rein dezeptiven Modus.

Interessanterweise kommen Menschen, die an einer rein sprachlichen Kontextbildung beteiligt sind, in Erklärungen häufig auf ihr *Vorstellungsvermögen* zu sprechen. Sie reden sogar davon, „innere Bilder“ zu haben, welche jene sprachliche Kontextbildung hervorgerufen hätten bzw. als Fundament der sprachlichen Kontextbildung dienten. Nun handelt es sich dabei, wie schon oben gesagt, nicht um etwas, das unter den Begriff ‚Bild‘ in dem Sinne, wie er bisher in dieser Abhandlung näher bestimmt wurde, fällt.

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 betrachtet; nur gibt es im aktuellen Verhaltenskontext keinen zugehörigen Bildträger. Oder es handelt sich um jemanden, der einer Fußball-Livereportage im Radio zuhört und sich die sportlichen Ereignisse ‚vor Augen führt‘: der sich also – für sich selbst – darstellt als jemand, der jenem Geschehen wahrnehmend beiwohnt und nicht bloß einen sprachlichen Bericht davon bekommt. Ihm scheint das Geschehen empirisch und nicht bloß logisch gegenwärtig.<sup>28</sup>

Es ist also die Situation der initialen Kontextbildung, in der sich zu befinden jene Person ‚vortäuscht‘, indem sie von ihrem Vorstellungsvermögen spricht. Das Vortäuschen einer initialen Kontextbildung fände dabei ebenfalls vor allem in einer verinnerlichten, nur für sich selbst vergegenwärtigten Form des Vorführens statt. Als Objekt, das den dezptiven Modus (die Täuschung) in diesem Falle vermittelt, träte allerdings die Person selbst in Erscheinung mit ihrem Körper und dem damit gegebenen Wahrnehmungs- und Ausdrucksvermögen; ein weiterer externer Bildträger wird nicht benötigt. So, wie zuvor ein in einer Zeichenhandlung benutztes Verhalten, ausgelöst von einem tatsächlich anwesenden Gegenstand, auf eine andere äußere Realität ausgerichtet wurde, 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 zugehörigen Bildträger *B'* – zu charakterisieren wäre. Und das ist nun tatsächlich eine mögliche Erklärung einer von der Anwesenheit eines wahrnehmungsnahen Zeichens unabhängigen, nicht-empirischen, d.h. einer nun rein logischen Form der Kontextbildung.<sup>29</sup>

Man schreibt einer Person (auch sich selbst) also dann ‚innere Bilder‘ oder den Gebrauch des (visuellen) Vorstellungsvermögens im Zusammenhang mit dem Gebrauch einer rein logischen (sprachlichen) Kontextbildung zu, wenn man auf eine Vergegenwärtigung der initialen Kontextbildung (also mittels eines wahrnehmungsnahen Zeichens) zur Erklärung der nicht durch einen täuschenden Gegenstand ausgelösten Kontextbildung verweisen möchte. So macht man aufmerksam auf die begriffsgenetische Rekonstruktion des Begriffsfeldes von Wesen, welche die Fähigkeit zum Gebrauch von Bildern, Sprache und sortalen Gegenständen besitzen. Es ist womöglich gerade der

28 Jörg R. J. SCHIRRA: Understanding Radio Broadcasts On Soccer. The Concept „Mental Image“ and Its Use in Spatial Reasoning. In: Klaus SACHS-HOMBACH (Hg.): Bilder im Geiste. Zur kognitiven und erkenntnistheoretischen Funktion piktorialer Repräsentationen. Amsterdam 1995, S. 107–136.

29 Diese Argumente finden sich ausführlicher in Jörg R. J. SCHIRRA, Klaus SACHS-HOMBACH: Fähigkeiten zum Bild- und Sprachgebrauch. In: Deutsche Zeitschrift für Philosophie 54 (2006), S. 887–905.

verinnerlichte (aber prinzipiell kommunizierbare) Bezug auf die Situation der initialen Kontextbildung, der auch beim Bildgebrauch stabilisierend gegen das Zurückfallen in den rein dezptiven Modus wirkt. Das würde auch erklären, warum von Kunsttheoretikern häufig von einem ‚inneren Bild‘ gesprochen wird, das als Voraussetzung für ein vom Künstler geschaffenes Werk zu gelten habe.

Dass, um zuletzt noch das eingangs erwähnte Beispiel aufzugreifen, die Schimpansin Moja tatsächlich das Bild eines Vogels gemalt und also Bildvermögen demonstriert hat, bleibt vor diesem Hintergrund eher fraglich. Zwar könnte es sich um die oben erwähnte Vorform piktorialer Kontextbildung handeln, wobei davon auszugehen wäre, dass die Farbmarken zunächst eher zufällig platziert wurden und erst im Anschluss die spontan-dezptive Reaktion auftrat und entsprechend kommunikativ eingebettet wurde (ASL-Zeichen für ‚Vogel‘). Doch fehlen weitere Belege dafür, dass Moja über Mechanismen verfügt, die zu einer sozial vermittelten Stabilisierung des immersiven Modus führen, so dass für sie weitere Kontexte nicht nur sporadisch, sondern systematisch neben die aktuelle Verhaltenssituation treten können. Ganz entsprechend bleibt auch unklar, in welcher Hinsicht Mojas Marken einem Vogel eigentlich visuell ähnlich sein sollen – unter anderem müssten dazu die Beziehung zu den Situationen, in denen Moja und ihre Gruppe auf tatsächlich anwesende Vögel reagieren, genauer beschrieben und analysiert werden.<sup>30</sup>

## 10. Résumé

Angeregt von Hans Jonas' phänomenologischer Reflexion zur anthropologischen *differentia specifica*, sind wir der begrifflichen Spur des *homo pictor* unter alternativer handlungstheoretischer Ausleuchtung gefolgt. Aus der in diesem Rahmen verwendeten Charakterisierung von Bildern als wahrnehmungsnahen Zeichen folgt, dass Gegenstände nur insofern Bilder sind, als sie in einem entsprechenden Verhalten, nämlich einem Kommunikationsverhalten (i. w. S.) von Bildnutzern eingesetzt werden. Durch eine solche Bestimmung wird eine gewisse Beziehung zu der argumentationstheoretischen Wende in der Philosophie des 20. Jahrhunderts deutlich, die als *linguistic turn* bekannt geworden ist: Steht der *linguistic turn* für die Erkenntnis, dass sich Begriffe (als die intersubjektiven Bezugspunkte für das Überprüfen der Geltung prädikativer Äußerungen) nicht unabhängig von Sprache bestimmen lassen, so kann der Ausdruck *pictorial turn*, der zunächst häufig nur zum Ausdruck bringen sollte, dass Umfang und Einfluss der Bildverwendung sich im Laufe der

30 Vgl. hierzu auch William NOBLE, Iain DAVIDSON: Human Evolution, Language, and Mind. A Psychological and Archaeological Inquiry. Cambridge 1996, S. 72ff.

letzten Dekaden drastisch verstärkt hat, ebenfalls argumentationstheoretisch verstanden werden: Auch die Bedeutung von Bildern entspringt nicht einer von ihrer Verwendung unabhängigen ‚objektiven‘ Relation zwischen Bildträger und Abgebildetem, sondern den Charakteristika der Verwendungssituation und damit bestimmter Handlungen.

Für den Bildgebrauch charakteristisch ergab sich im Anschluss eine bestimmte komplexe Verkopplung gewisser Handlungen (Zeichenhandlungen) mit bestimmten Handlungsdispositionen (im Zusammenhang mit der Handlung ‚Ähnlichkeit erkennen‘), die wir als immersiven Modus benannt haben. Wird mit einem Gegenstand in diesem Modus umgegangen, so wird er als wahrnehmungsnahes Zeichen verwendet. Wird die Ähnlichkeit dabei visuell erkannt, so handelt es sich um einen zentralen Fall von Bild. Die interessanteste Frage, nämlich die nach den genauen Voraussetzungen und Implikationen dieses speziellen Handlungsvermögens, die mithilfe einer begriffsgenetischen Rekonstruktion des entsprechenden Begriffsfeldes entlang der semiotischen und perzeptuellen Schienen auszuloten wäre, musste vorerst Forschungsprogramm bleiben. Lediglich der in anthropologischer Betrachtung vor allem interessierenden Frage nach der Beziehung zwischen Bildvermögen und Sprachvermögen wurde ein Stück weit nachgegangen, wobei die enge wechselseitige Abhängigkeit beider Vermögen deutlich wurde. Vieles spricht dafür, dass piktoriale Kontextbildung in der Tat als initialer Zugriff auf nicht-anwesende Situationen die assertorische Sprache – wie im Übrigen auch das Umgehen mit sortalen Gegenständen – begrifflich allererst ermöglicht und so direkt mit dem zentralen Kriterium des Menschseins im philosophischen Sinn verknüpft ist. Da umgekehrt auch der immersive Modus und damit das Bildvermögen erst dann stabil in Erscheinung treten können, wenn der Unterschied zum bloß dezeptiven Modus über das Sprachvermögen deutlich gemacht und der Bezug auf sortale Gegenstände ausdrücklich artikuliert werden kann, hängen beide Vermögen mit begrifflicher Notwendigkeit zusammen: Es kann keine Wesen geben, die nur über eines dieser Vermögen verfügen. Insofern also Sprachvermögen die zentrale anthropologische Differenz bildet, muss dies ebenso für das Bildvermögen gelten. Jonas ist also in dieser Hinsicht auch aus handlungstheoretischer Perspektive voll zuzustimmen.

Als Angelpunkt sowohl von Bildvermögen als auch Sprachvermögen, und somit für die anthropologische Differenz insgesamt, erweist sich die Fähigkeit zur Kontextbildung. Sie liefert den Zugang zu einem Lebensraum, der nicht mehr nur aus der eng begrenzten Raumzeit-Blase der jeweils aktuellen Verhaltenssituation besteht, sondern ein durch wechselseitig kontrollierte Regeln etabliertes System aus miteinander verbundenen Kontexten bildet, das es dem *homo pictor* erlaubt, prinzipiell ins Unendliche auszugreifen. Mit dieser Begabung wird – mit den Worten des französischen Paläoanthropologen André

Leroi-Gourhan – der Begriff einer *menschlichen Zeit* und eines *menschlichen Raums* allererst hervorgebracht.

„Der Mensch ist nur in dem Maße Mensch, wie er mit seinesgleichen zusammenkommt und sich mit den Symbolen seiner *raison d'être* umgibt. Nackt und bloß ist der Hohepriester wie der Vagabund nichts als der Kadaver eines höheren Säugetiers in einer Zeit und einem Raum ohne Bedeutung, denn sie sind nicht länger der Halt eines symbolisch menschlichen Systems. [...] Die menschliche Tatsache *par excellence* ist vielleicht weniger die Schöpfung des Werkzeugs als die Domestikation von Zeit und Raum, d. h. die Schöpfung einer menschlichen Zeit und eines menschlichen Raums.“<sup>31</sup>

### Auswahlbibliographie

- Hans BELTING: Bild-Anthropologie – Entwürfe für eine Bildwissenschaft. München 2001.
- Gottfried BOEHM (Hg.): Homo Pictor (Colloquium Rauricum, Bd. 7). München / Leipzig 2001.
- Merlin DONALD: Origins of the modern mind. Cambridge (MA) 1991.
- Hans JONAS: Die Freiheit des Bildens – *Homo pictor* und die *differentia* des Menschen. In: Zeitschrift für Philosophische Forschung 15 (1961), S. 161–176.
- André LEROI-GOURHAN: Hand und Wort. Die Evolution von Technik, Sprache und Kunst. Frankfurt am Main 1984.
- Steven MITHEN: The Prehistory of the Mind. A Search for the Origins of Art, Religion and Science. London 1996.
- William NOBLE, Ian DAVIDSON: Human Evolution, Language, and Mind – A Psychological and Archeological Inquiry. Cambridge (MA) 1996.
- Klaus SACHS-HOMBACH: Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft. Köln 2003.
- Jörg R. J. SCHIRRA, Klaus SACHS-HOMBACH: Fähigkeiten zum Bild- und Sprachgebrauch. In: Deutsche Zeitschrift für Philosophie 54 (2006), S. 887–905.
- Michael TOMASELLO: Die Ursprünge der menschlichen Kommunikation. Frankfurt am Main 2009.

31 André LEROI-GOURHAN: Hand und Wort. Frankfurt am Main 1984, S. 387.