

# ELiS\_e

[e'li:zə]

<Essener Linguistische Skripte\_elektronisch>

*Jörg R. J. Schirra &  
Klaus Sachs-Hombach*

**Bild und Wort.**

**Ein Vergleich aus bildwissenschaftlicher Sicht**

[elise@uni-due.de](mailto:elise@uni-due.de)

<http://www.elise.uni-essen.de>

**Jörg R. J. Schirra & Klaus Sachs-Hombach**<sup>1</sup>  
Otto-von-Guericke-Universität Magdeburg

## **Bild und Wort**

### **Ein Vergleich aus bildwissenschaftlicher Sicht**

Es hat eine lange Tradition, den Menschen als sprachbegabtes Tier zu charakterisieren. Dabei ist insbesondere an den Gebrauch prädikativer Satzstrukturen gedacht, also an Aussagen und davon abgeleitete Äußerungstypen (etwa entsprechende Frage- oder Imperativformen). 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? Eine solche Frage gehört ebenso gut in die allgemeine Sprachwissenschaft und Sprachphilosophie, wie sie der Bildwissenschaft eignet, die sich als von dem Fach Kunstgeschichte unabhängige Disziplin erst seit kurzem zu formieren beginnt (vgl. Sachs-Hombach 2003). Als ein erster Schritt zur Klärung aus bildwissenschaftlicher Perspektive sollen im Folgenden vor allem Gemeinsamkeiten von und Unterschiede zwischen dem Gebrauch von Sprache und dem Gebrauch von Bildern zusammengestellt werden. Doch erscheint es zunächst wichtig, eine genauere Bestimmung dessen, was mit „Bildwissenschaft“ gemeint ist, zu geben.

#### **1. Bildwissenschaftliche Fragestellungen**

Von Zeitgenossen wird die Bildwissenschaft, die sich im Grunde erst seit wenigen Jahren als eigenständige „Trans-Disziplin“ formiert, häufig in gewisser Konkurrenz zur bereits länger bestehenden Kunstgeschichte<sup>2</sup> gesehen. Daher lohnt es sich, zunächst mit dem Verhältnis dieser beiden Fächer auch zu klären, womit sich die Bildwissenschaft überhaupt auseinandersetzt. Sehen wir davon ab, dass die traditionelle Kunstgeschichte Bilder und andere Artefakte in der Regel in künstlerischem Zusammenhang, also in einer besonderen Verwendungsweise unter vielen, untersucht: Hiervon unabhängig ist wichtig, dass sie

<sup>1</sup> Zu den Autoren: <http://www.jrjs.de> & <http://isgnw.cs.uni-magdeburg.de/~ksh/homepage.html>

<sup>2</sup> Wir verwenden den Ausdruck „Kunstgeschichte“ hier stets im Sinne der akademischen Disziplin, die manchmal auch „Kunstgeschichtswissenschaft“ genannt wird, und nicht im ebenfalls möglichen Sinne von „faktische historische Entwicklung oder Abfolge von Kunststilen“.

spezifische Eigenarten von konkreten Bildwerken analysiert – und zwar vor einem oft nur impliziten Hintergrund, was Bildsein im Allgemeinen ausmacht. Damit ist durchaus eine indirekte Klärung des Allgemeinen durch die Untersuchung des Besonderen kompatibel, doch bleibt dieser Aspekt ein Nebenprodukt. Kunsthistoriker kann man (sofern sie sich nicht mit Skulpturen, Architektur oder anderen nicht im engeren Sinne bildhaften künstlerischen Artefakten befassen) durchaus auch charakterisieren als Forscher, die sich der wissenschaftlichen Betrachtung einzelner Bildwerke und ihrer Zusammenhänge vor allem (aber nicht nur) in historischer Entwicklung widmen. Daran ändert auch die reine Ausdehnung des Interesses von der ursprünglichen Beschränkung auf Kunstwerke hin zu allen visuell gestalteten Gebilden nichts wesentlich, soll es doch auch dabei gerade nur um „die Methoden der materialen Bestimmung, der historischen Zuschreibung und der semantischen Deutung“ jener einzelnen Gebilde gehen [Bredenkamp 2003, 56]. Kunstgeschichte bleibt so „Bildwissenschaft“.

Demgegenüber sollte von *Bildwissenschaft* im hier vertretenen Sinne dann die Rede sein, wenn sich das wissenschaftliche Interesse der Frage zuwendet, was es *grundsätzlich* bedeutet, mit Bildern (als solchen) umgehen zu können. Wie soll man es sich, zum Beispiel, vorstellen, dass sich eine solche merkwürdige Fähigkeit – oder genauer, dass sich Wesen mit einer solchen Fähigkeit – entwickelt haben? Und was folgt aus den so gewonnenen Charakterisierungen für die mit Bildern beschäftigten Einzelwissenschaften (insbesondere natürlich die Kunstgeschichte)? Die Bildwissenschaft versucht mithin zu klären, was Bildsein *allgemein* bedeutet. Damit klärt sie den von der Kunstgeschichte und anderen „Bildwissenschaften“ jeweils nur verwendeten und vorausgesetzten Hintergrund und wirkt so auch interdisziplinär als Vermittler. Das Untersuchen spezieller Einzelfälle ist dabei nicht unwichtig, dient aber vor allem der Verdeutlichung genereller Eigenschaften. Insbesondere gilt: Es stehen tatsächlich gar nicht einzelne Bilder im unmittelbaren Fokus des Interesses, sondern vielmehr die Fähigkeit, Bilder verwenden (d. h. erzeugen und rezipieren) zu können; und damit müssen von Bildwissenschaftlern, genau genommen, als Objekte ihrer Studien die Wesen, die über diese Eigenschaft verfügen, betrachtet werden. Noch präziser formuliert: Es geht nicht einmal um einzelne solcher Wesen, sondern um das ihnen allen Charakteristische, d. h. den *Begriff*,<sup>3</sup> den wir uns von Wesen mit der erwähnten Fähigkeit auf sinnvolle und rational kontrollierte Weise bilden können (und sollten).<sup>4</sup>

Dass die konkreten einzelnen Verwendungen eines bestimmten Bildes und die abstrakte allgemeine Fähigkeit zum Bildgebrauch aufs Engste miteinander verflochten sind, dass sie zwei Seiten ein und derselben Medaille bilden, das darf uns nicht verwundern: Der allgemeine Begriff eines Wesen, das Bilder gebrauchen kann, ist nichts anderes als der Komplex der abstrakten Regeln jenes Gebrauchs, die jeden empirischen Fall determinieren, in dem das Wesen diese Fähigkeit anwendet. Andererseits legt zugleich jede konkrete Anwendung

<sup>3</sup> Unter einem Begriff wird hier, wie in der Philosophie üblich, eine interindividuell kontrollierte Unterscheidungsfähigkeit verstanden (vgl. etwa Ros 1999).

<sup>4</sup> Ganz in diesem Sinn weitete Hans Belting seine ursprünglich kunstgeschichtlichen Untersuchungen seit einigen Jahren aus zu einer durchaus auch bildwissenschaftlich zu verstehenden *Bildanthropologie* (vgl. Belting 2001).

die nur abstrakt fassbaren und daher interpretationsbedürftigen Regeln weiter aus und erlaubt uns auf diese Weise überhaupt, uns dem Begriff wissenschaftlich zu widmen.

Damit besteht zwischen Bildwissenschaft und Kunstgeschichte ein ähnliches Verhältnis wie zwischen Literaturwissenschaft und Linguistik:<sup>5</sup> Die Sprachwissenschaft unterscheidet bekanntlich zwischen einzelnen Sprachzeichen, konkreten Sprachen (Systemen von Sprachgebräuchen) und der allgemeinen Sprachfähigkeit. Entsprechend dürfte es sinnvoll sein, zur Bestimmung des Forschungsgegenstandes der Bildwissenschaft zu differenzieren zwischen (1) bestimmten Bildern, (2) konkreten Bildgebräuchen und (3) der generellen Fähigkeit zum Bildgebrauch. Ein zentraler Gegenstand der Linguistik ist bekanntlich die *Sprachfähigkeit*; die Untersuchung einzelner Sprachen oder einzelner Sprachzeichen ist dafür vor allem Mittel zum Zweck.<sup>6</sup> Eine zentrale Fragestellung der Literaturwissenschaft betrifft hingegen insbesondere die Wechselwirkungen zwischen konkreter Form und konkretem Gehalt eines einzelnen literarischen Werkes. Dazu gehören z. B. auch Analysen rhetorisch-stilistischer Mittel wie Metaphern, Metonymien, Allegorien etc. Ganz ähnlich hat man sich in kunstgeschichtlichen Bild-Untersuchungen weniger um das eigentümliche Phänomen gekümmert, dass wir Menschen Wesen sind, die überhaupt mit etwas so Merkwürdigem wie Bildern umgehen können, sondern ganz wesentlich darum, wie bei einzelnen Bildern – gerade auch in historischem Entwicklungszusammenhang – die jeweiligen konkreten Bildformen und Bildgehalte miteinander korrespondieren. Auch hier treffen wir auf rhetorisch-stilistische Mittel wie Metapher, Allegorie usw.<sup>7</sup>

Für die Bildwissenschaft sollte hingegen – analog zur Linguistik – gelten: Es sind weniger die Unterschiede zwischen verschiedenen Bildverwendungen, die gesucht werden, sondern vor allem die Faktoren, die bei *jedem* einzelnen Gebrauch von Bildern eine Rolle spielen. Das wesentliche Ziel des Forschungsinteresses ist, wenn man es genauer formulieren möchte, nicht, wie der Name missverständlich andeutet, das Bild, sondern vielmehr die allgemeine Fähigkeit, Bilder verwenden zu können. Sowohl der konkrete, als Bild benutzte Gegenstand (der Bildträger bzw. das Medium) wie auch die abstrakten Funktionen der Bilder sind ausschließlich hinsichtlich dieses allgemeinen Vermögens eines Wesens zum Bildgebrauch bestimmt.

Aus dieser Nähe zwischen Linguistik und „Visualistik“, wie man die Bildwissenschaft analog auch nennen kann,<sup>8</sup> entspringt denn auch die Fragestellung, der wir uns mit dieser

<sup>5</sup> Hier wie dort bedingen und befruchten sich beide Bereiche auch wesentlich gegenseitig. – Die rein forschungspolitische Frage, ob Bildwissenschaft als Teil der Kunstgeschichte zu bewerten sei oder ob umgekehrt Kunstgeschichte besser als eine der an dem transdisziplinären Unternehmen ‚Bildwissenschaft‘ beteiligten Disziplinen zu sehen sein sollte, spielt auf einer ganz anderen, hier nicht weiter betrachteten Diskussionsebene und zielt letztlich vor allem darauf ab, wer die Kontrolle über entsprechende, insbesondere finanzielle Forschungsressourcen ausüben darf.

<sup>6</sup> Diese bilden ansonsten die Gegenstände der einzelnen Philologien wie Germanistik, Anglistik etc.

<sup>7</sup> In diesem Sinne wird man beispielsweise die Studien von Edgar Wind zu den Darstellungen der drei Grazien in der Renaissance verstehen dürfen (vgl. Wind 1984).

<sup>8</sup> Zumindest taucht diese Bezeichnung im Rahmen des Magdeburger Studiengangs *Computervisualistik* auf: Dort fasst der Titel „Allgemeine Visualistik“ die entsprechenden Lehrveranstaltungen zusammen (vgl. Schirra & Strothotte 1998).

Arbeit zuwenden: Welche begrifflichen Zusammenhänge bestehen zwischen der von der Bildwissenschaft untersuchten Fähigkeit, Bilder gebrauchen zu können, und der von der Linguistik behandelten Fähigkeit, Sprache gebrauchen zu können?

Schränken wir uns dazu aber zunächst etwas ein: Im Folgenden geht es, sofern nichts anderes erwähnt ist, stets um darstellende Bilder von bzw. um aussagenartige Sprachäußerungen zu raumzeitlich-materiellen konkreten Szenen. Dabei soll es allerdings keine Rolle spielen, ob es sich um reale oder fiktive Szenen handelt. In den Abschnitten 2 bis 5 werden auf einer allgemeinen Ebene entsprechende Sprach- und Bildverwendungen hinsichtlich verschiedener Aspekte miteinander verglichen. Abschnitt 6 dient dazu, den Zusammenhang mit Strukturbildern sowie Aussagen über Abstraktes anzudeuten. Der abschließende siebente Abschnitt beginnt mit einer Zusammenfassung und verweist als Ausblick auf das Programm einer umfassenderen begriffsgenetischen Betrachtung, mit deren Hilfe es möglich sein sollte, die gefundenen begrifflichen Gemeinsamkeiten und Unterschiede zwischen Bild- und Sprachvermögen auf systematische Weise zu begründen und damit zugleich eine Fundierung der Bildwissenschaft zu geben.

## 2. Eine gemeinsame Basis: Die Gebrauchssituation als Zeichen

Auf den ersten Blick gibt es sowohl eine Reihe wichtiger Ähnlichkeiten zwischen Bild- und Sprachgebrauch als auch bedeutende Unterschiede. Es scheint zum Beispiel ein zentraler Verwendungszusammenhang von Bildern zu sein, dass eine Person allein und ganz für sich ein Bild betrachtet.<sup>9</sup> Hingegen wird allgemein eine prinzipiell dialogische Situation der Verwendung von Sprache zugrunde gelegt:<sup>10</sup> Jemand gibt jemand anderem etwas zu verstehen (oder intendiert das mit der Sprechhandlung doch zumindest). Da sprachliche Zeichen notwendigerweise nicht einfach in der Natur vorzufinden sind, sondern von einem Sprecher produziert sein müssen, ergibt sich automatisch, dass ihr Gebrauch wesentlich eine Interaktion zwischen zwei Partnern – genauer: eine Kommunikation – ist.<sup>11</sup> Bei Bildern gibt es demgegenüber, jedenfalls auf den ersten Blick, durchaus Fälle, die natürlich entstanden zu sein scheinen: Man denke etwa an Schattenwürfe

<sup>9</sup> Für die Kunstgeschichte hatte sich dieser Verwendungsmodus geradezu als paradigmatisch herausgebildet, wurde aber in jüngerer Zeit zunehmend aufgeweicht (vgl. Bogen 2005). Allerdings fordern einige zeitgenössische Kunstwissenschaftler nach wie vor, dass jeder, der wissenschaftlich-vernünftig über Bilder reden wolle, sich – ähnlich einem Flugzeug-Kapitän, der eine Mindestanzahl an Flugstunden absolviert haben muss, um seiner Aufgabe gewachsen zu sein – zunächst jahrelang viele Stunden täglich auf solch private Weise mit Bildern beschäftigt, sich in sie versenkt haben müsse, bevor er (oder sie) in der Lage sei, mitreden zu können.

<sup>10</sup> Das gilt jedenfalls spätestens seit dem Zusammenbruch der neuzeitlichen Bewusstseinsphilosophie und dem darauf folgenden *linguistic turn*: Was wir als Bedeutung sprachlicher Zeichen verstehen kann nicht als etwas, was bereits sprachunabhängig vorliegt und erst nachträglich mit sprachlichen Mitteln kommunizierbar gemacht wird, konzipiert werden. Es ergibt sich erst mit und in sprachlicher Interaktion (vgl. Dummett 1992).

<sup>11</sup> Um die interessante Ausnahme, die durch die Produkte der sogenannten sprach-generierenden Computersysteme gegeben zu sein scheint, können wir uns an dieser Stelle nicht kümmern (vgl. Schirra 2005, Kap. 4).

oder spiegelnde Flächen. In der Tat werden Vorkommen solcher „natürlichen Bilder“, wie sie auch genannt werden, bereits in den antiken Bildtheorien als Auslöser für die Fähigkeit des Bildgebrauchs genannt (vgl. Plinius 1977, 23; siehe dazu auch Scholz 2000, 623).

Mehr als dem Gespräch ähnelt das einsame Betrachten eines Bildes, wenn man einen Vergleich mit dem Gebrauch sprachlicher Zeichen anstrebt, dann vielleicht dem Lesen, das ja auch in der Regel einer, der möglichst dabei nicht gestört werden möchte, für sich allein betreibt. Doch genau besehen handelt es sich auch beim Lesen um eine kommunikative Aktivität, nur dass hier nicht ohne Weiteres klar ist, wer eigentlich die Rolle des Senders einnimmt. Demgemäß haben uns die Literaturwissenschaftler darauf aufmerksam gemacht, dass man literarische Werke, die uns hier als ein typischer Fall von gelesener Sprache dienen mögen, prinzipiell in (mindestens) zwei sich in der Konzeption des Senders unterscheidenden Weisen rezipieren kann: Zum einen gibt es einen historischen Urheber des Textes, den man (in seiner jeweiligen Situation) als den Sender betrachten kann: „Was will der Autor uns (oder genauer: den von ihm damals anvisierten Lesern) damit sagen?“ Zum anderen aber kann man sich selbst als Sender sehen und versuchen herauszufinden, was mit dem betrachteten Text jetzt, in der aktuellen Situation, zu verstehen gegeben werden kann: „Was kann einer meiner Zeitgenossen (oder ich selbst in der Rolle eines solchen Zeitgenossen) mir damit hier und heute sagen?“ In beiden Konzeptionen liegt also tatsächlich eine direkt vom Sprechen abgeleitete dialogische Verwendungssituation vor, bei der allerdings einer der Interaktionspartner internalisiert ist und nur in der Vorstellung als Gegenüber auftritt.<sup>12</sup> Viele andere scheinbar monologische Gebrauchsformen von Sprache lassen sich auf ähnliche Weise kommunikativ verstehen.

Aber ähnelt das Verinnerlichen des Kommunikationspartners nicht auch dem einsamen Gebrauch eines Bildes? Ist es nicht so, dass ich mir, wenn ich ein Spiegelbild betrachte, selbst (in der Rolle eines anderen, der von jener Stelle aus schauen könnte) etwas zeige? Und mache ich mich nicht selbst, wenn ich in einer Galerie ein Bild studiere – es mir sozusagen selbst präsentiere –, auf dieses oder jenes aufmerksam? Das mag etwa ein Aspekt des Bildgehaltes oder ein Stilmerkmal sein, eine physische Eigenschaft des Bildträgers, wie etwa Alterungsspuren oder sogar eine vorgestellte andere Verwendungssituation (die Präsentation des Bildes in einer typischen dialogischen Situation einem anderen gegenüber). Es ist durchaus sehr plausibel, auf diese Weise auch einsame Bildverwendungen auf kommunikative Handlungen zurückzuführen: Nicht der Bild-„Monolog“ ist dann die Stan-

<sup>12</sup> Auch für den Autor wird entsprechend eine internalisierte dialogische Verwendungssituation mit vorgestellten Lesern vorausgesetzt. Umgekehrt tritt ebenso der historische Autor, da er in der Regel einem Leser nicht persönlich bekannt sein dürfte, meist als nur vorgestellte historische Person in den internalisierten dialogischen Situationen auf. Darauf, dass das Internalisieren der Interaktionspartner eine wesentliche Voraussetzung für die Fähigkeit zu bewusster Kommunikation ist, hat im Übrigen bereits G. H. Mead hingewiesen (vgl. Mead 1934).

dardgebrauchssituation, sondern der Bild-„Dialog“.<sup>13</sup> In dieser Hinsicht besteht also starke Ähnlichkeit zwischen Bildgebrauch und Sprachgebrauch.

Gehen wir also im Weiteren davon aus, dass man sowohl beim Bildgebrauch wie beim Sprachgebrauch immer davon sprechen kann, dass es einen (gegebenenfalls interiorisierten und vielleicht sogar nur imaginierten) Sender und einen (ebensolchen) Empfänger gibt, die kommunikativ miteinander interagieren (zumindest in der Vorstellung). Genau in diesem noch recht unspezifischen Sinn sprechen wir im Folgenden von einer *Zeichenhandlung*.

### 3. Wahrnehmungsnähe

Ein bedeutender Unterschied zwischen Bild- und Sprachgebrauch ergibt sich, wenn wir die Wahrnehmungsnähe, also die Besonderheit der visuellen Wahrnehmungskompetenzen für den Bildgebrauch, ins Auge fassen. Denn in der Tat spricht vieles dafür, dass zumindest beim Gebrauch bestimmter (und sogar als besonders charakteristisch eingeschätzter) Bilder auf solche Wahrnehmungsfähigkeiten zurückgegriffen wird, die eigentlich zum Sehen ganz anderer Gegenstände und Szenen dienen: Und das sind gerade jene Gegenstände und Szenen, von denen man sagt, sie seien in jenen Bildern „abgebildet“. Dass man oft davon spricht, dass Bilder dem Abgebildeten ähnlich seien, drückt genau diese Wahrnehmungsnähe aus. Eine vergleichbare Kopplung mit Wahrnehmungskompetenzen lässt sich für den Sprachgebrauch nicht ohne Weiteres erkennen.

Ein aus dieser Beobachtung abgeleitetes Verständnis von Bildern als einer speziellen Sorte *wahrnehmungsnaher Zeichen* (nämlich den *visuellen* wahrnehmungsnahen Zeichen) spiegelt sich in vier unterschiedlichen Reflexionsmodi, die beim Umgehen mit einem als Bild benutzbaren Gegenstand (d. h. einem potentiellen Bildträger) auftreten können.<sup>14</sup> Zunächst einmal kann man sich einfach täuschen und den (potentiellen) Bildträger mit dem Abgebildeten spontan verwechseln: Wem ist das nicht schon passiert, wenn er sich einem täuschend echt gemalten Bild, einem *Trompe l’Œil*, gegenüber befand. Wir sprechen hier vom *dezeptiven* Modus oder Modus der Täuschung, den im Übrigen auch höhere Tiere durchaus einnehmen können (vgl. Abb. 1).

Andererseits kann man den Bildträger als Zeichen verwenden, wobei aber noch nicht an ein wahrnehmungsnahes Zeichen gedacht sei. Man erkennt also die kommunikative Situation, erkennt, dass da ein Gegenstand anwesend ist, mit dessen Hilfe ein Sender einem Empfänger etwas in der Regel nicht zugleich Anwesendes vergegenwärtigen – repräsentieren – möchte. Das ist der *symbolische* Modus oder Modus der Zeichenverwendung, der sich insbesondere dadurch auszeichnet, 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. 2).

<sup>13</sup> Auch „natürliche Bilder“ sind also erst dann und nur dann tatsächlich Bilder, wenn sie in eine entsprechende (verdeckte oder offene) dialogische Situation eingebunden sind. Ansonsten ist eine optisch reflektierende Fläche eben auch nur das: eine optisch reflektierende Fläche.

<sup>14</sup> Diese Reflexionsstufen treten bei allen Arten wahrnehmungsnaher Zeichen entsprechend auf.

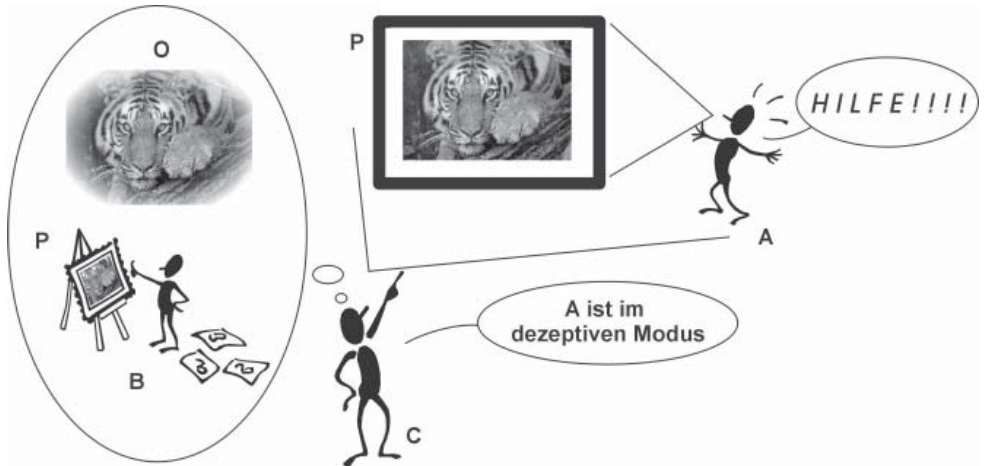


Abbildung 1: Der dezeptive Modus

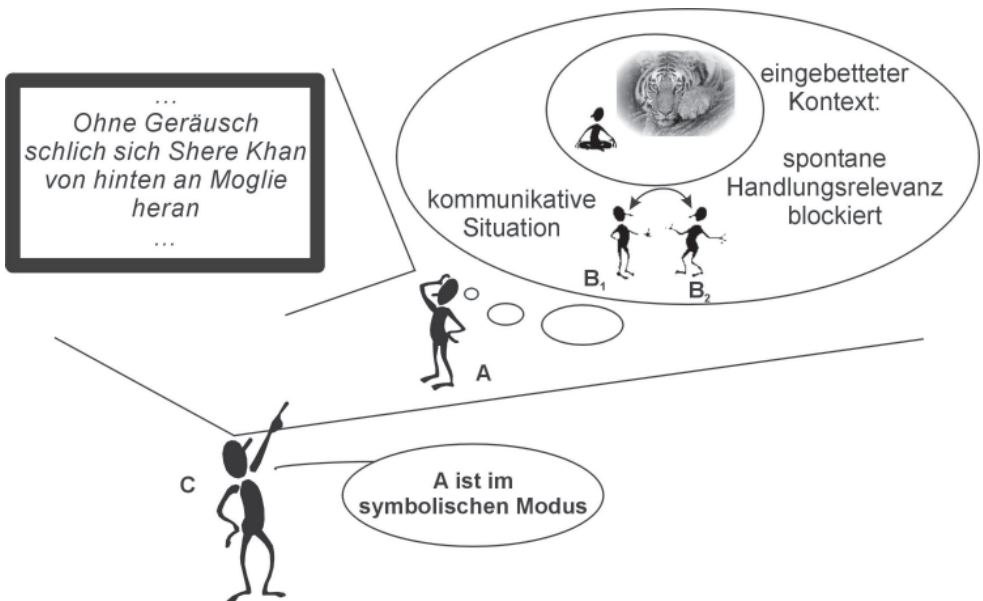


Abbildung 2: Der symbolische Modus

Für uns besonders interessant wird es, wenn die beiden erwähnten Modi auf systematische Weise zusammenspielen: wenn man also die Täuschung sowohl erlebt als auch durchschaut und so als Grundlage für eine Zeichenverwendung nimmt. Das ist der *immersive* Modus oder Modus der symbolischen Immersion. Er ist die Grundlage des Gebrauchs wahrneh-

mungsnaher Zeichen. Denn als Zeichen erfordern sie natürlich den symbolischen Modus. Andererseits soll der Zeichenträger bei wahrnehmungsnahen Zeichen dem Repräsentierten ähnlich sein. Er muss also eine mehr oder weniger starke spontane Verwechslungsreaktion auslösen, was dem dezptiven Modus entspricht. Diese spontane Reaktion tritt allerdings – wegen der Einbettung in den Modus der Zeichenverwendung – in der Regel nicht nach außen, sondern wird nur als das wahrscheinlich gemeinte Repräsentierte wirksam und bewusst (vgl. Abb. 3).

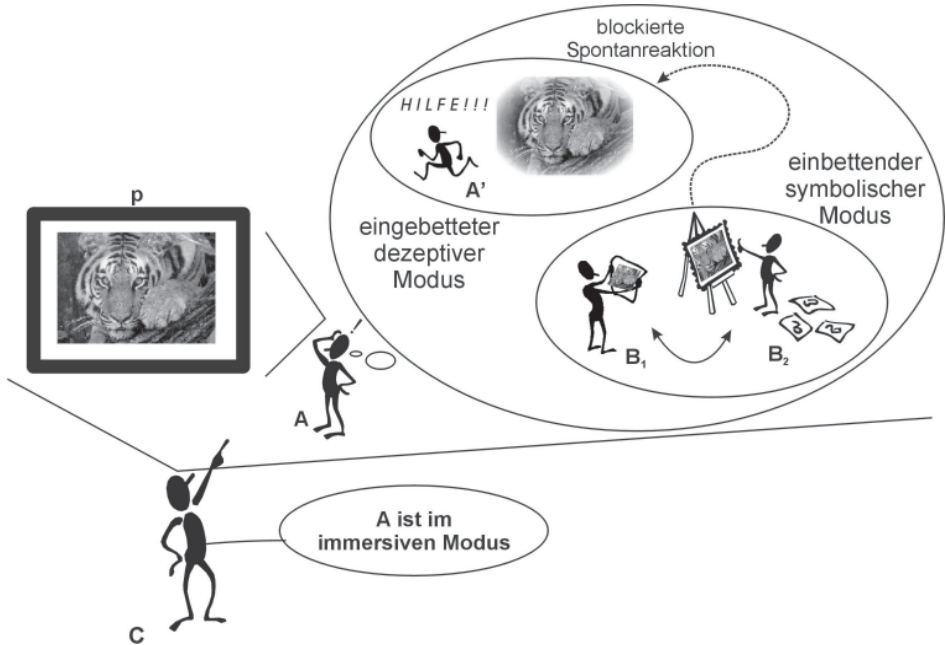


Abbildung 3: Der immersive Modus

Schließlich kann ein Zeichen auch im *reflexiven* Modus verwendet werden, wobei allerdings in der Regel nicht die eigentliche Bedeutung des Zeichens, also etwa das in einem Bild Abgebildete, gemeint ist, sondern auf Aspekte der entsprechenden Zeichenverwendung selbst exemplarisch aufmerksam gemacht wird. Viele Bilder der Kunst sollten in diesem Modus verwendet werden.<sup>15</sup>

<sup>15</sup> Natürlich muss man zugeben, dass nicht allen Bildern ohne Weiteres ein Abgebildetes zugeordnet werden kann. An dieser Stelle sei nur erwähnt, dass die Bilder ohne Abgebildetes einen Gebrauch im reflexiven Modus erzwingen: Das Fehlen des an sich erwarteten Abgebildeten löst als Fehlkommunikation nämlich entsprechende „Fehleranalysen“ aus, durch die dann Aufmerksamkeit auf die Bildkommunikation selbst gelenkt wird. In solchen Fällen sprechen wir von „reflexiven Bildern“ (im Unterschied zu darstellenden und logischen (auch: Struktur-)Bildern). Sie spielen allerdings in dieser Abhandlung keine weitere Rolle. Abgebildetes muss im Übrigen auch nicht immer sichtbar sein. Der negative Fall zeichnet gerade Strukturbilder aus (siehe dazu auch Abschnitt 6).

Vergleichen wir dies mit dem Gebrauch von Sprache, so sind offensichtlich nur der symbolische und der reflexive Modus denkbar. Der symbolische Modus versteht sich von selbst, denn er entspricht ja gerade dem normalen Gebrauch von Sprache. Auch für den reflexiven Modus fallen sofort Beispiele ein, denn der Gebrauch eines Zitats gehört hier ebenso her wie die Verwendung von Beispielsätzen in linguistischen Abhandlungen. Dass man aber einem sprachlichen Zeichen (oder genauer seinem Zeichenträger) im dezeptiven Modus gegenüberreten könnte, dass man den Zeichenträger also mit dem so Bezeichneten spontan verwechseln könnte, ist bei gesunden Sprechern kaum vorstellbar und stellt, wenn es denn doch auftritt, eine große Ausnahme dar. Daraus folgt dann sofort, dass es einen immersiven Modus für Sprache nicht (ohne Weiteres) geben kann, da dieser Modus als Kombination von symbolischem und dezeptivem Modus bestimmt ist.

Fassen wir zusammen: Die Nutzung im immersiven Modus kann geradezu als konstituierend für den Begriff „wahrnehmungsnahes Zeichen“ betrachtet werden. Der immersive Modus ist zugleich etwas, was dem Sprachgebrauch fehlt. Speziell auf die visuelle Wahrnehmung bezogen, führt er damit zu einer genaueren handlungstheoretischen Charakterisierung des Bildgebrauchs.

#### **4. Spezifische Aufgliederungen**

Eine besondere Eigenschaft sprachlicher Handlungen besteht darin, dass sie in mehrfacher Hinsicht aus Teilen aufgebaut sind. Dabei spielt die auch jedem Laien bekannte syntaktische Aufteilung nicht einmal die wichtigste Rolle.

##### **4.1 Illokutionäre Funktion, Prädikation und Nomination**

Wesentlich relevanter sind die funktional-pragmatische Aufteilung in illokutionäre Funktion und propositionalen Gehalt einerseits, und die Aufteilung des propositionalen Gehalts in Prädikation und eine oder mehrere Nominationen andererseits. Die Aufteilung einer sprachlichen Zeichenhandlung in illokutionäre Funktion und propositionalen Gehalt trägt der Tatsache Rechnung, dass man, indem man Sprache benutzt, in der Regel auch noch etwas anderes tut als eben nur zu sprechen [Austin 1979]: Man warnt oder verspricht, fragt oder befiehlt, behauptet oder bezweifelt, um nur einige Beispiele für illokutionäre Funktionen zu nennen.

Von diesem Aspekt, der sich auf die mit der Sprachhandlung vollzogene Interaktion bezieht, ist der Sachbezug zu unterscheiden: das, wovon man warnt, was man verspricht, erfragt oder befiehlt.<sup>16</sup> Der den Sachbezug vermittelnde propositionale Gehalt besteht wiederum aus zwei Arten von Teilhandlungen: 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 auf welche einzelnen Gegenstände) er sich dabei beziehen möchte. Mit einer sol-

<sup>16</sup> Allerdings gibt es auch sprachliche Handlungen ohne propositionalen Gehalt, etwa das Grüßen.

chen Teilzeichenhandlung alleine lässt sich keine vollständige Kommunikationshandlung vollziehen. Prädikation und Nomination gelten – wie es in der Terminologie Freges (1892) heißt – jeweils als „ungesättigt“.

Lassen sich beim Bildgebrauch ähnliche Teilhandlungen identifizieren? Sicher kann man mit Bildern auch illokutionäre Akte vollziehen: z. B. vor einem bissigen Hund warnen oder darauf hinweisen, dass es sich bei einer Tür um diejenige der Damentoilette handelt. Allerdings wird allgemein bezweifelt, dass alle illokutionären Funktionen, die mit Sprache möglich sind, auch allein durch das Präsentieren eines passend gewählten Bildes ausgeführt werden können. Wie soll man sich etwa einen piktorialen Akt des Bezweifeln vorstellen, oder ein bildhaftes Versprechen? Aber auch das Umgekehrte mag wahr und die sprachliche Handlung in einigen Hinsichten dem Bildgebrauch unterlegen sein.

Der propositionale Gehalt stellt uns beim Bildgebrauch sogar vor größere Probleme, als man zunächst annehmen mag: Auf den ersten Blick scheinen Bilder ohne Weiteres beide der Propositionen konstituierenden Teilfunktionen übernehmen zu können. Die Standardfunktion eines Passfotos etwa könnte prädikatorisch verstanden werden: „Dieser Mensch *sieht so (!) aus*“, wobei das Foto eine sehr komplizierte visuelle Unterscheidungsgewohnheit anspricht, die linguistisch durch die deiktische Partikel „so“ samt Zeigehandlung als (Teil der) Prädikation eingebunden wird, und die Nomination implizit in der Situation klar ist (nämlich auf denjenigen, der den Pass präsentiert hat). Umgekehrt empfinden wir es als nicht ungewöhnlich, wenn jemand das Bild einer großen, rot angestrichenen Hängebrücke mit zwei auffällig gestalteten Pfeilern präsentiert und dazu knapp bemerkt: „Ist 1936 gebaut worden.“ Der Bildzeichenakt übernimmt also, wie es scheint, die Rolle der Nomination; hier gilt ebenso wie bei den rein sprachlichen Nominationsen, dass auch der empfangende Diskurspartner jenes Objekt bereits als gemeinsam bekannt verstehen muss, da er ansonsten nicht wüsste, worauf er die Prädikation des Baujahres, die als zuvor noch nicht gemeinsam bekannt gilt, überhaupt beziehen soll.

In beiden Fällen übernimmt anscheinend die bildhafte Zeichenhandlung – das Präsentieren des Bildes – eine Teilhandlung einer Aussage, deren sprachliche Komponente ganz im Frege’schen Sinn ungesättigt ist, ohne Bild also nicht wirklich verständlich wäre. Zwar kommen auch ansonsten Aussagen häufig vor, deren sprachliche Form unvollständig (elliptisch) ist und etwa keine Nomination zu enthalten scheint: „Ganz schön frech!“ Doch sind in diesen Fällen die scheinbar fehlenden Teile lediglich implizit kommuniziert worden: Sie ergeben sich u. a. durch die Aufmerksamkeitsfokussierung auf unmittelbar vorher erwähnte oder in der Situation auffällig präsente Gegenstände.

Die erwähnten Beispiele mit den Bildern sind ganz anders gelagert, da hier, wie es scheint, die verbal fehlenden Teilhandlungen durch Bildzeichenakte explizit zu einer vollständigen Aussage gesättigt werden. Daher drängt sich einerseits der Verdacht auf, dass auch Bildzeichenhandlungen ungesättigt sind. Doch dann wären Kommunikationsakte mit Bildern, die keine Teilhandlungen komplexerer Gesamtzeichenhandlungen sind, nicht vorstellbar. Zum anderen fällt auf, dass Bildzeichenhandlungen nicht *per se* oder auf sonstwie offensichtliche Weise einem der beiden Aussage-Teilhandlungen eindeutig

zugeordnet werden können. Die Verwendungsweise des Bildes wird in solchen Fällen offenbar vor allem von den ansonsten auftretenden sprachlichen Aussage-Teilhandlungen und deren Ergänzungsbedürftigkeit induziert.

## **4.2 Figur, Grund und Medium**

Obwohl also vieles dafür spricht, dass eine Aufgliederung für Bildpräsentationen entsprechend der für Sprache relevanten Einteilung der propositionalen Akte nicht ohne Weiteres vorliegt, gibt es doch auch für Bilder ein gleichermaßen spezifisches wie fundamentales Phänomen pragmatischer Aufgliederung, nämlich die Unterscheidung in Figur und Grund. Diese Differenzierung ist tatsächlich nicht an wahrnehmungsnahe Zeichen gebunden, sondern betrifft zunächst das Wahrnehmen ganz allgemein. Wenn wir wahrnehmen, werden – von unserer je aktuellen mentalen Befindlichkeit geleitet – perzeptuelle Merkmale organisiert und so jeweils entweder als Teil einer als relevant eingestuftes Figur interpretiert oder als Teil des Grundes, in den jede wahrgenommene Figur notwendig eingebettet ist: Einer Figur wird im Weiteren deutlich mehr Aufmerksamkeit geschenkt. Diese Aufgliederung ist variabel nicht nur in dem Sinne, dass unsere Aufmerksamkeit nicht ständig auf alle Gegenstände der Umgebung fokussiert ist, sich vielmehr von einem zum anderen bewegt, so dass jeweils was zunächst Figur ist zu einem Teil des Hintergrundes der folgenden Figur wird. Veränderlich ist die Differenzierung in Figur und Grund insbesondere auch in dem Sinne, dass eine völlige Uminterpretation der aktuellen Verhaltenssituation möglich bleibt. Das geschieht zum Beispiel auch dann, wenn wir eine Spiegelfläche, die wir erst unbewusst als eine andere Person gesehen hatten (dezeptiver Modus), als spiegelnde Oberfläche eines Brettes erkennen.

Auch im Bild ist es eine auf diese Weise – also wahrnehmend – gewonnene Figur, die uns jeweils besonders interessiert, und die dabei vor einen Grund tritt, der sie gleichzeitig einhüllt und positioniert. Während sich Prädikation und Nomination nur unzureichend von ihrer sprachlichen Herkunft auf Bilder übertragen lassen, ist es wesentlich leichter, umgekehrt die Figur-Grund-Unterscheidung von ihrem wahrnehmungstheoretischen und bildbezogenen Ursprung auf die Teilhandlungen des propositionalen Gehalts von Sprache zu beziehen. Denn die Prädikation funktioniert ja ähnlich einer Figur: Sie hebt eine als relevant erachtete Unterscheidungsgewohnheit, die in dem durch Nomination gegebenen, bereits als bekannt vorausgesetzten „Grund“ verankert sein muss, in den Fokus der – in diesem Falle *gemeinsamen* – Aufmerksamkeit.<sup>17</sup>

Allerdings handelt es sich auch bei der Zuordnung von Grund und Figur im Bild immer um eine von mehreren dem Bild inhärenten Möglichkeiten. Einige Kippfiguren (vgl. Abb. 4) spielen mit dieser Mehrdeutigkeit: Sie sind so angelegt, dass es spontan zu drastischen Figur-Grund-Wechseln kommt, die nicht miteinander kompatibel sind. Allgemein gilt: Sind

<sup>17</sup> Statt der Begrifflichkeit „Nomination und Prädikation“ wird in der Prager Schule der Linguistik auch das recht eng verwandte Begriffspaar „Thema“ und „Rhema“ betrachtet, das insbesondere die Differenz der beiden Teile hinsichtlich dessen, was aktuell gemeinsam bekannt bzw. für den Empfänger neu ist, hervorhebt.

mehrere Objekte abgebildet, so kann jedes einzelne davon zur Figur vor den Rest des Bildes treten. Ist ein einzelner, aber komplexer Gegenstand abgebildet, so können auch jeweils Teile oder Aspekte von ihm als Figur auftreten – vor jeweils sich entsprechend ergebenden Hintergründen, die diese Figur relational erfassen. Im oben schon als Beispiel erwähnten Photo der *Golden Gate*-Brücke etwa ist zwar die Brücke eine herausragende Möglichkeit für die Wahl der Figur; aber man kann mit dem Bild ebenso gut die Aufmerksamkeit eines Empfängers auf die Meerenge, über die sich die Brücke spannt, richten, oder auf das abgebildete Segelboot, das gerade die Bucht verlässt; auf die Stadt, die ebenfalls (teilweise) abgebildet sein mag, oder ein einzelnes ihrer im Bild sichtbaren Gebäude; auf den einen Brückenpfeiler, der im Wasser steht, oder den großen grauen Fleck, der anzeigt, wo der charakteristische rote Schutzanstrich der Brücke erneuert wird. In jedem Fall kommt eine andere Figur-Grund-Differenzierung zum Tragen, bei welcher der jeweilige Grund notwendig ist zur Verankerung dessen, was an der Figur letztlich als relevant gesehen wird.

Im Gegensatz zu der den jeweiligen propositionalen Gehalt überhaupt erst determinierenden Aufteilung in eine festgelegte Prädikation und eine (oder mehrere) ebenfalls eindeutig definierte Nomination(en) ist die verwandt erscheinende inhaltliche Aufteilung in Figur und Grund bei Bildern gerade nicht fest und eindeutig vorgegeben. So, wie beim Wahrnehmen die Umwelt eine Bühne bietet, die jeweils in Relevantes und den das Relevante erfassenden Rest erst zergliedert wird, bieten Bilder vor allem *Möglichkeiten* für unterschiedliche Figur-Grund-Aufteilungen.<sup>18</sup>

Diese Eigenschaft, Figur-Grund-Aufteilungen anzubieten, und das Verhältnis dieser Eigenschaft zu der manifesten Aufteilung in Prädikation und Nomination im propositionalen Gehalt sprachlicher Äußerungen werden noch genauer zu untersuchen sein. Um eine einfache und nicht ganz unmotivierte Ausdrucksweise verfügbar zu haben, schlagen wir vor, das noch nicht realisierte Potential einer Entität zu Figur-Grund-Differenzierungen – an eine Terminologie aus der Gestaltpsychologie anknüpfend – als „Medium“ zu bezeichnen.<sup>19</sup>

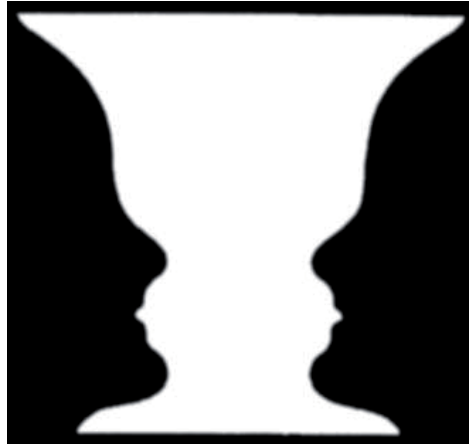


Abbildung 4: Kippfigur „Rubin'scher Becher“

<sup>18</sup> Allerdings bietet sich bei Bildern – im Gegensatz zur Wahrnehmungssituation – auch an, bildrhetorische Mittel einzusetzen, indem etwa durch entsprechende Variationen im Darstellungsstil bestimmte Aspekte so markiert werden, dass sie eher als Figur zu verstehen sind (vgl. Sachs-Hombach & Schirra 2002).

<sup>19</sup> Vgl. Heider 1927 sowie die Gegenüberstellung von „figuralem und medialem Ich“ in Bischof 1998. Es mag misslich erscheinen, den ohnehin bereits ausgesprochen vieldeutigen Ausdruck

Figur und Grund differenzieren sich also stets innerhalb eines Mediums, das die Möglichkeit dieser Differenzierung bereitstellt. Sie sind auf dieses Medium bezogen, das zugleich für alternative Figur-Grund-Aufteilungen verfügbar bleibt. Bilder liefern uns Medien in diesem Sinne. Die propositionalen Gehalte sprachlicher Äußerungen beruhen hingegen auf einer Art vollzogener Figur-Grund-Differenzierung; doch bleibt die Frage vorerst offen, welches Medium es ist, auf das diese sich bezieht.

## 5. Kontextbildung: Empirische und logische Vergegenwärtigung

Sprachliche Äußerungen mit propositionalem Gehalt werden, im Gegensatz zu einfacheren Zeichenhandlungen, vor allem dann verwendet, wenn wir über Situationen kommunizieren, die mit der Äußerungssituation nichts unmittelbar zu tun haben (vgl. etwa Tugendhat 1976). Dabei beziehen wir uns in der Regel mit den Nominationen auf eine besondere Art von raum-zeitlich-materiellen Gegenständen, die in der Philosophie häufig als *sortale* Gegenstände bezeichnet werden.<sup>20</sup>

Wir nehmen es üblicherweise als gegeben hin, dass unsere Welt (neben anderem) vor allem aus individuellen Gegenständen besteht: aus Tischen und Stühlen, Bäumen, Katzen, Autos und Häusern 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 verändert haben (z. B. Raupe zu Schmetterling). Nur wegen der sortalen Individuation kann festgestellt werden, ob es sich zu zwei verschiedenen Zeitpunkten um *denselben* Gegenstand handelt oder um zwei nur *gleichartige* Gegenstände. Um die Klärung einer solchen Identität geht es etwa, wenn ein Gericht festzustellen versucht, ob es sich bei dem Dolch, der jetzt auf dem Asservatentisch 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. 5). Auch die möglichen Transformationen, die Gegenstände der betrachteten Sorte verwandeln können, ohne deren Identität zu ändern (etwa das Rosten des Messers, Verschmutzung, das Umbiegen der Messerspitze bei unsachgemäßem Gebrauch), sind als Teil des sortalen Begriffs determiniert.

---

„Medium“ hier ins Spiel zu bringen. Gut motiviert wird die Wahl durch den Zusammenhang von Medien in diesem Sinne und Verhaltenssituationen ganz allgemein: Die Umwelt eines Wesens ist, ohne Weiteres, ebenfalls ungegliedert. Sie wird erst gemäß der jeweiligen Motivationen des Wesens gegliedert, wenn dieses ihr wahrnehmend und sich verhaltend gegenübertritt, oder wohl besser: in sie – als Medium – eintaucht.

<sup>20</sup> Genau genommen spricht man von *sortalen Begriffen*. „Sortale Gegenstände“ ist hier eine abkürzende Sprechweise für „Gegenstände, die unter einen sortalen Begriff fallen“.

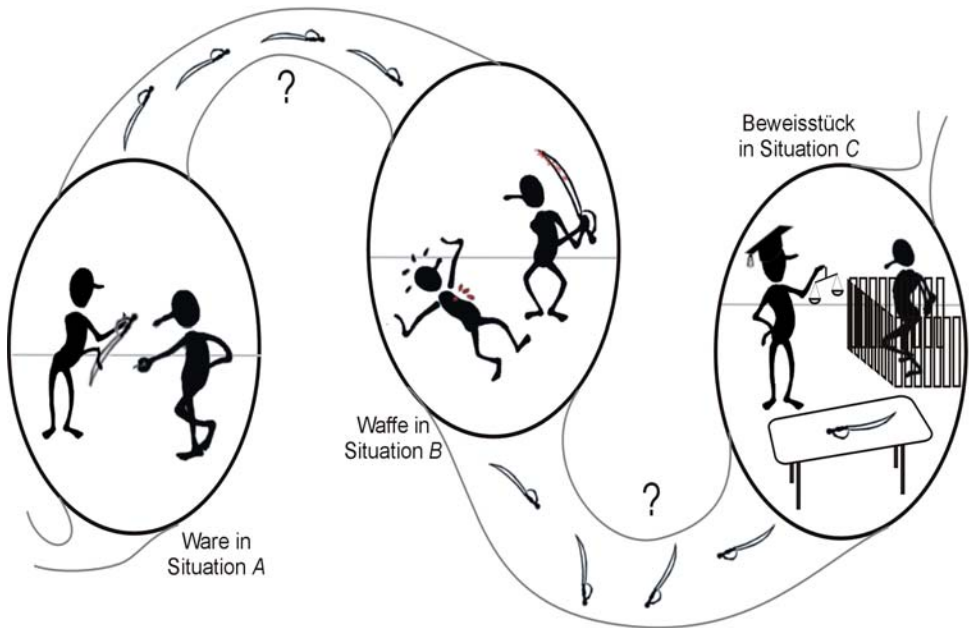


Abbildung 5: Identität und Lebenslinie eines sortalen Gegenstands

Es ist wichtig, sich hier klarzumachen, dass sortale Gegenstände nie isoliert vorkommen: Es macht immer nur Sinn, von ihnen als etwas zu reden, was als Figur vor einem Hintergrund in Erscheinung tritt und in mehr als einem Verhaltenskontext existiert. Mit „Kontext“ soll hier zunächst jede endliche, strukturierte Menge von intentionalen sortalen Gegenständen bezeichnet werden: also eine Reihe von miteinander in Beziehung stehenden Gegenständen, insoweit jemand etwas über sie weiß (oder sie wahrnimmt).

Da sich Propositionen in der Regel nicht auf den aktuellen Verhaltenskontext beziehen, müssen wir neben Prädikation und Nomination notwendigerweise eine weitere Teilhandlung berücksichtigen, durch die der jeweils gemeinte Kontext für das Gegenüber spezifiziert wird. In Anlehnung an den Ausdruck „*space builder*“ des Linguisten Gilles Fauconnier (1985) sei diese Funktion „Kontextbildung“ genannt. Sprachliche Orts- oder Zeitangaben können der Kontextbildung dienen. Aber auch hypothetische oder fiktive Kontexte können verbal gebildet werden, etwa durch Verweise auf Texte literarischer Fiktion: „In Uwe Johnsons Roman ‚Jahrestage‘ hat Gesine eine Tochter namens Marie.“ Wen der Sprecher mit „Marie“ oder „Gesine“ meint, ist nur relativ zu dem angegebenen literarisch-fiktiven Kontext verständlich. Mit der sprachlichen Kontextbildung wird so neben der aktuellen Äußerungssituation ein weiterer situativer Kontext eröffnet, so dass sich die Sprecher im

Weiteren aus der Distanz auf die den Kontext aufspannenden sortalen Gegenstände beziehen können.<sup>21</sup>

Der aktuelle situative Kontext spielt sicher eine ausgezeichnete Rolle, denn nur bei den Aussagen, die sich auf ihn beziehen, greift die referentielle Verankerung von Nominati- on und Prädikation unmittelbar. Allein hier können die senso-motorischen Komponenten der verwendeten Begriffe – also die entsprechenden Unterscheidungsgewohnheiten – ohne Weiteres ausgespielt werden. Der situative Kontext eignet sich deshalb auch als Standardan- nahme: Es erscheint plausibel und „ökonomisch“, den situativen Kontext der Dialogpartner bei der Äußerung – d. h. die Äußerungssituation – zugrunde zu legen, sofern keine andere Kontextbildung ersichtlich ist.

Kontextbildungen, die wie das Satzadverbial „in Prag“ auf Orte verweisen, geben ge- wissermaßen eine Methode an, wie der gemeinte Kontext in den aktuellen situativen Kon- text überführt werden könnte, um dann die eigentliche referentielle Verankerung des pro- positionalen Gehalts der Äußerung durchzuführen. Um Aussagen referentiell zu verankern, muss man demnach stets zweierlei beachten: Man muss erstens verstehen, wie man die mit der Nomination verbundenen senso-motorischen Routinen zur Identifikation der gemeinten Gegenstände „positioniert“ (indem man den durch Kontextbildung angegeben Kontext zur aktuellen Verhaltenssituation macht); und man muss zweitens wissen, wie die mit der Prädikation assoziierten senso-motorischen Tests in der dann aktuellen Situation durchzuführen sind (wie die gemeinte Unterscheidungsgewohnheit zu praktizieren ist).

So weit deckt sich der Begriff des Kontextes mit der oben angegebenen Bestimmung als endlicher, strukturierter Menge von intentionalen sortalen Gegenständen. Allerdings kön- nen wir Kontexte offenbar ebenso gut als verallgemeinerte Aktivitätszusammenhänge, als *Verhaltenssituationen* etwa für senso-motorische Testroutinen bestimmen. Mit Wittgenstein sollte man Verhaltenssituationen nun nicht einfach als Gegenstandsmengen auffassen, viel- mehr als Systeme von Sachverhalten: das, was der Fall ist *dann und dort*. Was wir mit „Ver- haltenssituationen“ bezeichnen, bildet jeweils ein Interpretationsangebot, eröffnet ein Po- tential zu vielen Reaktionen oder Auslegungen in Form von Aussagen (vgl. Fellmann 2000, 27ff.). Denn Verhaltenssituationen sind prinzipiell für viele verschiedene Figur/Grund-Un- terscheidungen und damit für Perspektivwechsel offen. Die Charakterisierung als Gegen- standsmenge ist daher verkürzt: Sie ist gewissermaßen die *figurale* Sicht auf einen Kontext. Ihr sollte eine *mediale* Sicht als interpretations-offener Verhaltenskontext zur Seite treten.

Kontexte bilden somit einerseits – insofern nämlich auch die aktuelle Verhaltenssituati- on ein Kontext ist bzw. jeder andere Kontext (zumindest prinzipiell) zur aktuellen Verhal- tenssituation gemacht werden kann – das Medium, das die Wahrnehmung in gegenständ- liche Figuren und Hintergründe zu trennen vermag und das dann als Menge intentionaler Gegenstände erscheint. Andererseits bilden sie – insofern sie durch Propositionen erzeugt und in Propositionen repräsentiert werden können – auch das Medium für die ganz anders geartete Differenzierung in Prädikation und Nomination: als System von Sachverhalten, die auf unterschiedliche Weise als propositionale Gehalte kommuniziert und dabei in die

<sup>21</sup> Auch Beziehungen zwischen den beiden (oder weiteren durch Kontextbildung eingeführten) Kontexten können sprachlich artikuliert werden, wozu insbesondere Identitätssätze dienen.

Figur einer als noch nicht gemeinsam bekannt angesehenen Unterscheidungsgewohnheit und einen Hintergrund von als bereits bekannt vorausgesetzten (d. h. in schon gemeinsam gewusste Sachverhalte eingebunden) Gegenständen unterschieden werden kann.

\* \* \*

Offenbar stellt der nicht-reflexive Gebrauch eines darstellenden Bildes ebenfalls einen zusätzlichen situativen Kontext zur Verfügung, der als Ausgangspunkt für Nominationen und Prädikationen dienen kann und in aller Regel auch dient. Der Gebrauch des Bildes liefert den Kommunikationspartnern jedenfalls genau den Kontext, der für die Aussagen über den Bildinhalt (d. h. das Abgebildete) benötigt wird. Ähnlich wie im weiter oben angegebenen Beispiel die Erwähnung eines Romans den Gesprächspartnern einen ganz bestimmten (fiktiven) Kontext für die anschließenden Nominationen und Prädikationen eröffnet, versucht jemand durch die Präsentation eines Bildes, die Aufmerksamkeit des Empfängers auf eine reale oder fiktive Wahrnehmungssituation zu richten, die in den meisten Fällen nicht mit der tatsächlichen Situation der beiden übereinstimmt.<sup>22</sup>

Dass in der Tat jeder andere Gebrauch als Bild auf der Verwendung zur Kontextbildung beruht erscheint dann als vielversprechende Hypothese: Denn es ist stets zunächst das scheinbare sinnliche Vergegenwärtigen einer meist nicht zugleich anwesenden Situation, das speziellere Gebräuche überhaupt erst möglich macht. Beim Einsatz als Passbild beispielsweise evoziert das Vorzeigen des Bildes einen Kontext, der es erlaubt, eine andere Erscheinung des zu identifizierenden sortalen Gegenstands (genauer, die Person, die jenen sortalen Gegenstand als Leib hat) als Referenz für die Identitätsprüfung zu nutzen.

Die Auffassung, dass Bildgebrauch im Kern Kontextbildung sei, bedeutet übrigens keineswegs, dass eine endliche Menge von Aussagen einem Bild äquivalent wäre. Überhaupt kann keine bestimmte Menge von Aussagen den Zeichengehalt eines Bildes vollkommen bestimmen – noch übrigens irgendeiner anderen direkt situativ oder durch Kontextbildung gegebenen Verhaltenssituation. Verstanden in seinem medialen Aspekt schränkt der spezifizierte Kontext zwar die möglichen Nominationen und Prädikationen ein; er determiniert sie aber nicht vollständig. Vielmehr gehört es, wie bereits erwähnt, zur Funktion der Kontexte für das propositionale Zeichenverhalten, dass sie gegenüber den möglichen durch Nominationen und Prädikationen gegebenen Interpretationen offen bleiben müssen.

Die rein sprachliche Kontextbildung muss in der Regel ohne eine auf Wahrnehmung gestützte *empirische* Verifikation der jeweiligen Behauptung auskommen. Lediglich bei Aussagen über die aktuelle Äußerungssituation kann eine referentielle Verankerung direkt stattfinden; ansonsten muss erst – sofern das überhaupt möglich ist – der durch Kontextbildung spezifizierte Kontext konkret aufgesucht werden. Die Kontextbildung mit Hilfe von Bildern bietet hingegen wegen des eingebetteten dezeptiven Modus zumindest innerhalb gewisser Grenzen (etwa der Sinnesmodalität) durchaus auch die Möglichkeit, die entsprechenden Nominationen und ihr Zusammenspiel mit der erwähnten Prädikation unmittelbar auf Wahrnehmung gestützt – also wie im aktuellen Äußerungskontext – zu verankern.

<sup>22</sup> Wie in Abschnitt 2 erwähnt, kann es sich durchaus auch um die internalisierten Rollen in einer nur vorgestellten Kommunikationssituation handeln.

Um das in der Aussage behauptete Zutreffen bestimmter Begriffe auf die angegebenen Einzelgegenstände empirisch überprüfen zu können, müsste man die aktuelle Verhaltenssituation verlassen und sich in den spezifizierten Kontext begeben. In diesem Sinne stehen sprachlich eingeführte Kontexte gewissermaßen deutlich getrennt nebeneinander: Man kann nicht zugleich in mehreren davon referentielle Verankerungen einlösen. Will oder kann ich den erwähnten Kontext nicht zum aktuellen Kontext machen (etwa bei Fiktion), bleibt sogar nur übrig, durch logische Schlüsse weitere Aussagen aus den verwendeten Begriffen abzuleiten und auf Konsistenz mit dem bereits über jenen Kontext Gewussten zu untersuchen. Doch gelangt man so natürlich nie zu einer Situation, in der das Gelten der Behauptung tatsächlich empirisch nachgeprüft werden könnte.

Bei der Kontextbildung durch das Präsentieren eines Bildes wird hingegen eine zusätzliche Verhaltenssituation heraufbeschworen, die im dezptiven 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 vielmehr teilweise miteinander „verschmolzen“. Mit anderen Worten: Verbale Kontextbildung kann Sachverhalte nur *logisch vergegenwärtigen*, während piktoriale Kontextbildung sie *empirisch vergegenwärtigen* kann.

## **6. Kurze Bemerkung zu Strukturbildern und abstrakter Sprache**

In den vorausgegangenen Überlegungen standen vor allem bildliche Darstellungen räumlich-materieller Szenen einerseits und der sprachliche Zugriff auf raumzeitliche Gegebenheiten mittels Aussagen andererseits im Interessenfokus. Bildliche und sprachliche Darstellungen von raumzeitlich Konkretem sind zwar für ihr jeweiliges Gebiet zentrale Phänomene, doch wird damit bei Weitem nicht schon alles abgedeckt. Insbesondere muss sich eine umfassende Ausarbeitung auch den Verwendungen von Strukturbildern und von auf Abstraktes bezogener Sprache widmen. Während für Letztere nicht ohne Weiteres deutlich ist, auf welche Weise hier die empirischen Geltungsprüfungen erfolgen können, bleibt bei Strukturbildern zunächst unklar, auf welche Weise hier Wahrnehmungsnahe als Basis für das empirische Vergegenwärtigen zu bestimmen wäre. Denn Strukturbilder (oder logische Bilder) sind bestimmt als solche Bilder, deren Abgebildetes eigentlich gar nicht gesehen werden kann (z. B. graphische Darstellung zu Wahlwanderbewegungen) oder doch mindestens einen nicht-visuellen Aspekt aufweist (z. B. Wärmebild eines Hauses).

Vieles spricht dafür, dass beide Übergänge miteinander zusammenhängen, da metaphorische Beziehungen zwischen Begriffsfeldern, wie sie etwa in der Kognitiven Linguistik untersucht werden (vgl. etwa Lakoff & Johnson 1980, Johnson 1987), dabei eine wichtige Rolle spielen dürften. Genauer gesagt sind partielle Strukturübertragungen von dem Feld der sortalen Gegenstände auf ein Begriffsfeld für die betrachtete abstrakte Sphäre zu berücksichtigen, durch die es letztlich, wie die Kognitive Linguistik behauptet, überhaupt möglich wird, über jene abstrakten Gegenstände und ihre Eigenschaften und Beziehungen

zu reden; nämlich so über sie zu reden – aber eben auch: sie bildlich so darzustellen –, als wären sie sortale Gegenstände in einem Kontext mit entsprechenden visuell erkennbaren Parametern und Relationen (vgl. Schirra 2005, 4.4.4). Auch für Strukturbilder wäre also die Funktion der Kontextbildung zentral.

Doch sollen diese Überlegungen hier aus Platzgründen nicht weiter verfolgt werden.

## 7. Ausblick: Ein begriffsgenetisches Forschungsprogramm

Fassen wir kurz zusammen: Für den Sprach- wie den Bildgebrauch haben wir eine dialogisch-kommunikative Grundsituation vorausgesetzt. Eine Besonderheit des Bildgebrauchs haben wir dann in dem ausgezeichneten Bezug zu den (visuellen) Wahrnehmungskompetenzen gesehen. Die Unverzichtbarkeit dieses Bezuges für ein angemessenes Verständnis des Bildvermögens, auf die wir mit der Rede von wahrnehmungsnahen Zeichen hinweisen wollen, ist handlungstheoretisch durch einen spezifischen Rezeptionsmodus bestimmt, nämlich den immersiven Modus, der sich nicht ohne Weiteres beim Sprachgebrauch ausmachen lässt. Umgekehrt haben wir als Charakteristikum des Sprachgebrauchs eine Gliederung in komplementäre ungesättigte Teilhandlungen ausgemacht, die beim Bildgebrauch höchstens in speziellen, abgeleiteten Verwendungsweisen auftreten. Es deutete sich an, dass die wahrnehmungstheoretische Figur-Grund-Relation auch einer möglichen nominatorischen oder prädikatorischen Verwendung zugrunde liegt. Der Bildgebrauch verweist dabei auf die noch undifferenzierte Basis des Figur-Grund-Gegensatzes: das „Medium“, das auf verschiedene Weise aufgegliedert werden kann und mit den sortalen Gegenstandsbegriffen eng zusammenhängt. Bildgebrauch, so die daran anschließende These, ist im Kern Kontextbildung, mit der, im Gegensatz zur sprachlichen Kontextbildung, nicht-anwesende Verhaltenssituationen empirisch vergegenwärtigt werden. Diese These bedarf offensichtlich weiterer begrifflicher Klärung.

Nicht zuletzt zum Zweck einer solchen begrifflichen Klärung sollte die Bildwissenschaft genauer untersuchen, worin die logisch-begrifflichen Voraussetzungen bestehen, die wir jeweils unterstellen müssen, wenn wir von Wesen mit der Fähigkeit zum Bildgebrauch bzw. zum Sprachgebrauch reden. Und mehr noch: Sie sollte untersuchen, wie wir diese Voraussetzungen begründen können. Daraus lässt sich, so unsere Annahme, tatsächlich ein präziseres Forschungsprogramm für die Bildwissenschaft ableiten, das hier skizziert werden soll.

Zu diesem Zweck sei zunächst an einige methodologische Erkenntnisse der philosophischen Argumentationstheorie erinnert. Sie unterscheidet zwischen Begriffserläuterungen und Begründungsversuchen für Begriffserläuterungen (vgl. Ros 1999). 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. Nun werden allerdings in der Regel jene zur Erläuterung verwendeten Begriffe bei weiterer Rückfrage oft selbst wiederum unter anderem mit dem zuerst erläuterten Begriff erklärt. Systeme aus sich wechselseitig bestimmenden Begriffen nennt man auch ein *Begriffsfeld*. Solche zyklischen Ab-

hängigkeiten lassen sich meist lösen, indem eine kleine Menge von Grundbegriffen eines Begriffsfeldes festgelegt wird, auf die jeder andere Begriff des Systems logisch zurückgeführt werden kann. Die Grundbegriffe selbst sind dann allerdings 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 als durch eine systematische Kombination von zunächst als voneinander unabhängigen und von allen Parteien akzeptierten Begriffsfeldern eingeführt zu betrachten, 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 kontingenter Weise aufweisen, kommen unter dem kombinierten Begriffsfeld Instanzen vor, die Eigenschaften aus beiden elementarerer Begriffsfeldern auf systematische Weise gekoppelt zeigen. Dabei legt das Schema der Begriffsfeldkombination zusammen mit den internen Regeln der elementarerer Felder die Regeln fest, die für begriffsfeld-interne (logische) Begriffserläuterungen im kombinierten Feld gelten: eine Begründung der Axiome selbst sozusagen.

Die begriffsgenetische Betrachtung baut darauf, dass sich zweierlei zeigen lässt: (1) In keinem der elementarerer 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 elementarerer Begriffsfelder akzeptieren, das Schema der Kombination annehmen und überhaupt Interesse daran haben, gemeinsame Begriffe der betrachteten Art zu etablieren.

\* \* \*

Uns interessieren die Zusammenhänge zwischen den Begriffsfeldern für Wesen mit der Fähigkeit zum Bildgebrauch bzw. zum Sprachgebrauch. In den vorangehenden Betrachtungen haben wir im Wesentlichen einige begriffliche Erklärungen zusammengetragen und damit versucht, die internen Strukturen der jeweiligen Begriffsfelder etwas näher zu bestimmen. Dabei wissen wir noch nicht einmal, ob es sich dabei um Zusammenhänge in einem einzigen Begriffsfeld handelt (also Bildvermögen und Sprachvermögen nur gemeinsam auftreten können), oder ob etwa Bildgebrauch zu einem der für den Sprachgebrauch begriffsgenetisch konstitutiven Begriffsfelder gehört (oder umgekehrt). Dieser Frage

können wir nur durch entsprechende begriffsgenetische Betrachtungen zu Leibe rücken. Zugleich erhalten wir mit der so gewonnenen Begründung ein zusätzliches Kohärenzkriterium für die internen Bestimmungen des (jeweiligen) Feldes.

Aus den oben vorgetragenen Überlegungen zu Bild- und Sprachgebrauch ergeben sich für die geforderten begriffsgenetischen Betrachtungen erste Anhaltspunkte. Zeichengebrauch kann, so hatten wir gesehen, als gemeinsame Basis für beide Vermögen dienen. Daher müssen wir uns auf verschiedene Stufen der Komplexität der Begriffe von Wesen, die Zeichen gebrauchen können, einlassen. Andererseits können Wahrnehmungsnähe und das Verhältnis von Medium und Figur-Grund-Differenzierung als Ziel einer spezifischen Differenz verwendet werden. Wir müssen uns demnach auch unterschiedlich 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 tatsächlich in der Ethologie wie auch in der Sprachphilosophie vorzufinden (vgl. etwa auch Ros 2005).

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. Das damit gegebene minimale Feld mit einem Begriff für bildbenutzende Wesen müsste genau die wesentlichen Bestimmungsstücke versammeln. Von ihm ausgehend könnten – gegebenenfalls – komplexere Felder, die höhere Stufen des Bildvermögens charakterisieren, abgeleitet und so klar unterschieden werden. Auch sollte damit die Beziehung zu jener Stufe geklärt werden können, auf der der Gebrauch propositionaler Sprache möglich wird – eine Stufe, die ohnehin Teil der „semiotischen“ Stufenfolge sein muss. So wäre dann klar, ob das Sprachvermögen eine unumgängliche Voraussetzung des Bildvermögens ist oder umgekehrt; oder ob alternativ beide sich wechselseitig bedingen, so dass von Wesen, die nur über eines der beiden Vermögen verfügen, nicht die (vernünftige) Rede sein kann.

Es dürfte offensichtlich sein, dass die detaillierte Ausführung dieses Planes, der letztlich auf eine fundamentale Klärung der die Bildwissenschaft charakterisierenden Fragestellung abzielt, den Rahmen dieser Abhandlung bei Weitem sprengen würde. Immerhin hoffen wir, mit dieser kurzen Skizzierung nicht nur ein begriffsgenetisches Forschungsprogramm für die allgemeine Bildwissenschaft entworfen, sondern auch einen gangbaren Weg für den Dialog mit der Linguistik gewiesen zu haben.

\* \* \*

Wir möchten all denen danken, die uns bei der Arbeit an dieser Abhandlung ermutigt und konstruktiv unterstützt haben.

## Literatur

- Austin, John L. (1979): *Zur Theorie der Sprechakte (How to do things with words)*, Stuttgart: Reclam.
- Belting, Hans (2001): *Bild-Anthropologie – Entwürfe für eine Bildwissenschaft*, München: Fink.
- Bischof, Norbert (1998): *Das Kraftfeld der Mythen*. München: Piper.
- Bogen, Steffen (2005): Kunstgeschichte/Kunstwissenschaft, in: Klaus Sachs-Hombach (Hg.): *Bildwissenschaft. Disziplinen, Themen, Methoden*, Frankfurt/M.: Suhrkamp, 52–67.
- Bredenkamp, Horst (2003): Bildwissenschaft, in: *Metzler Lexikon Kunstwissenschaft*, Stuttgart: Metzler.
- Dummett, Michael (1992): *Ursprünge der analytischen Philosophie*, Frankfurt/M.: Suhrkamp.
- Fauconnier, Gilles (1985): *Mental Spaces: Aspects of Meaning Construction in Natural Language*, Cambridge: Cambridge Univ. Press.
- Fellmann, Ferdinand (2000): Bedeutung als Formproblem – Aspekte einer realistischen Bildsemantik. In: Sachs-Hombach, K. & Rehkämper, K. (Hg.): *Vom Realismus der Bilder: Interdisziplinäre Forschungen zur Semantik bildhafter Darstellungsformen*, Magdeburg: Scriptorum, 17–40.
- Heider, Fritz (1927): Ding und Medium, Berlin: *Symposium. Philosophische Zeitschrift für Forschung und Aussprache* 1, S. 109.
- Johnson, Mark (1987): *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: Univ. of Chicago Press.
- Lakoff, George & Johnson, Mark (1980): *Metaphors We Live By*. Chicago: University of Chicago Press.
- Mead, George Herbert (1934): *Mind, Self, and Society*, Chicago: University of Chicago Press.
- Plinius Secundus d. Ä. (1977): *Naturkunde (Naturalis Historiae)*, Lat.-dt., Buch XXXV: Farben, Malerei, Plastik, hg. und übersetzt von Roderich König in Zusammenarbeit mit Gerhard Winkler, 2., überarbeitete Auflage, Düsseldorf / Zürich: Artemis und Winkler.
- Ros, Arno (1999): Was ist Philosophie?, in: Raatzsch, Richard (Hg.): *Philosophieren über Philosophie*, Leipzig: Leipziger Universitätsverlag, 36–58.
- Ros, Arno (2005): *Materie und Geist. Eine philosophische Untersuchung*, Paderborn: Mentis.
- Sachs-Hombach, Klaus (2003): *Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft*, Köln: Herbert von Halem.
- Sachs-Hombach, Klaus & Schirra, Jörg R.J. (2002): Selecting Styles For Tele-Rendering. Toward a rhetoric in computational visualistics, in: *Proc. 2<sup>nd</sup> International Symposium on Smart Graphics 2002*, Hawthorne, NY: ACM Press, 102–106.
- Schirra, Jörg R.J. (2005): *Foundation of Computational Visualistics*, Wiesbaden: DUV.
- Schirra, Jörg R. J. & Strothotte, Thomas (1998): Von Bildern und neuen Ingenieuren: Aspekte des Studienganges Computervisualistik. In: A. Dreß & G. Jäger (Hg.): *Visualisierung in Kunst und Mathematik: Technik - Bilder - Visionen*. Braunschweig / Wiesbaden: Vieweg, 189–205.

- Scholz, Oliver R. (2000): Bild, in: Barck, Karlheinz u. a. (Hg.): *Ästhetische Grundbegriffe*. Historisches Wörterbuch in sieben Bänden, Bd. 1, Stuttgart: Metzler, 618–669.
- Tugendhat, Ernst (1976): *Vorlesungen zur Einführung in die sprachanalytische Philosophie*, Frankfurt/M.: Suhrkamp.
- Wind, Edgar (1984): *Heidnische Mysterien in der Renaissance*, 3. Aufl., Frankfurt/M.: Suhrkamp.

*Abbildung 4: aus Fisher, Gerald H. (1967): Measuring ambiguity. American J. Psychology 80, 541–557.*