

MARCEL LEMMES / STEPHAN PACKARD /  
KLAUS SACHS-HOMBACH

## Bilder im Aufbruch: Herausforderungen der Bildwissenschaft

In den 1990er-Jahren erlebten Bilder einen bemerkenswerten Aufschwung in Gesellschaft und Wissenschaft. In der langen Reihe der Turns wurde nun auch die Wende zum Bild ausrufen und zugleich mit zahlreichen Namen versehen. Neben dem *pictorial turn* (MITCHELL 1992) gab es – etwas früher bereits – den *imagic turn* (FELLMANN 1991) und – etwas später – den *iconic turn* (BOEHM 1994), schließlich ergänzt durch den *visualistic turn* (SACHS-HOMBACH 2003). Bereits einige Zeit zuvor war die Rede von der Bilderflut gebräuchlich geworden, die sich mit dem Siegeszug der technisch erzeugten Bilder eingestellt hatte, also vor allem mit Fotografie, Film und Fernsehen. Während der Verweis auf die Bilderflut einen eher kritischen Unterton zu den Bildern lieferte, waren insbesondere die kunsthistorischen Bemühungen zum *iconic turn* um eine Aufwertung der Bilder bemüht. Dies galt insbesondere in Relation zum *linguistic turn*, der mitunter als Fortsetzung der langen Geschichte der Bildabwertungen gesehen wurde und der die aktuelle Bildforschung zu sehr in die Nähe semiotischer oder linguistischer Modelle gebracht hätte. Der zeitgleiche Verweis auf Gefahren der Bilddominanz und die empathische Betonung stärker zu würdigender Besonderheiten der Bilder im Verhältnis zur Sprache entspricht ganz der ambivalenten Einschätzungen von Bildphänomenen, die sich wie ein roter Faden seit je nicht nur durch die christlich-abendländische Kulturgeschichte zieht.

Aktuell scheint die Geschichte der Bilder (und allgemeiner: der visuellen Medien) in eine neue Phase zu treten und weitere Ambivalenzen zu Tage zu fördern. Das betrifft zum einen die enorm beschleunigte Distribution

von Bildern in den sozialen Medien, zum anderen betrifft es die mittels KI eröffneten Möglichkeiten der Bildgenerierung, deren Effekte sich bisher noch gar nicht absehen lassen. Insbesondere haben sich die technischen Möglichkeiten der Bildmanipulation in einem Maße erhöht, dass sie die uns bisher bekannten Verfahren der Bildbearbeitung (ewtwa im Rahmen der digitalen Fotografie oder mithilfe von Bildbearbeitungsprogrammen wie Photoshop) als harmlose Spielerei und die ehemals übertrieben wirkenden Bilddystopien bei Vilém Flusser (1985) oder Jean Baudrillard (1994) als aktuelle Gesellschaftsdiagnosen erscheinen lassen.

Die Chancen und Gefahren, die sich im Schnittfeld von sozialen Medien und KI-generierten Bildern ergeben, lassen sich noch kaum überblicken. Entsprechend stellen diese zukünftig gesellschaftlich immer stärker wirkenden Phänomene auch für die Bildforschung eine besondere Herausforderung dar. Sie waren Gegenstand der interdisziplinären Fachkonferenz »Die Zukunft der Bilder – Herausforderungen der Bildwissenschaft«, die vom 8. bis 10. Februar 2024 an der Eberhard Karls Universität Tübingen stattfand. Bei der Planung und Durchführung haben wir erheblich von der Unterstützung durch das dortige Institut für Medienwissenschaft und die Universität profitieren können, wofür wir uns an dieser Stelle herzlich bedanken. Unser Dank gilt natürlich insbesondere auch allen Referentinnen und Referenten.

Der vorliegende Band versammelt eine Auswahl der Konferenzbeiträge. Mit der Organisation der Konferenz verfolgten die Herausgeber einerseits das inhaltliche Ziel einer kritischen Bestandsaufnahme neuer Bildformen. Andererseits sollte in institutioneller Ausrichtung insbesondere denjenigen Wissenschaftler:innen ein Forum zum wissenschaftlichen Austausch geboten werden, die sich seit den 1990er-Jahren in verschiedenen Konstellationen der Ausgestaltung einer interdisziplinären Bildwissenschaft verpflichtet fühlen. Dies betrifft etwa das »Zentrum für interdisziplinäre Bildwissenschaft«, das hieraus entstandene »Netzwerk der Bildphilosophie« oder die verschiedensten hiermit assoziierten Zusammenhänge, in denen gemeinsame Forschungen und Publikationen zum Bildthema entstanden waren. Der Einladung zum Austausch waren spontan viele gefolgt, sodass die gemeinsame Reflexion auf die Zukunft gleichermaßen der Bildkultur wie der Bildwissenschaft in üblich interdisziplinärer und stimulierender Weise ermöglicht wurde.

## Literatur

- BAUDRILLARD, JEAN: *Simulacra and Simulation*. Ann Arbor, MI [University of Michigan Press] 1994
- BOEHM, GOTTFRIED: Die Wiederkehr der Bilder. In: GOTTFRIED BOEHM (Hrsg.): *Was ist ein Bild?* München [Fink] 1994, S. 11–38
- FLUSSER, VILÉM: *Ins Universum der technischen Bilder*. Göttingen [European Photography] 1985
- FELLMANN, FERDINAND: *Symbolischer Pragmatismus. Hermeneutik nach Dilthey*. Hamburg [Rowohlt] 1991
- MITCHELL, WILLIAM J. THOMAS: The Pictorial Turn. In: *Artforum*, 30 (7), 1992, S. 89–94
- SACHS-HOMBACH, KLAUS: *Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft*. Köln [Herbert von Halem] 2003

JOACHIM KNAPE

## Probleme der Bildtheorie und Bildrhetorik

### 1. Bildwissenschaft und Bildtheorie

Als der seinerzeitige Magdeburger Privatdozent Klaus Sachs-Hombach im Jahr 2005 seine beiden Sammelbände zum Thema *Bildwissenschaft* publizierte, veranstaltete der Karlsruher Kunsthistoriker Hans Belting (1935–2023) im selben Jahr eine Konferenz zu *Bilderfragen* mit dem Untertitel *Die Bildwissenschaften* [Plural] *im Aufbruch* (BELTING 2007). Dieser Untertitel war eine Art Kampfansage an Sachs-Hombach, der schon seit einiger Zeit das Projekt einer ›Bildwissenschaft‹ [Singular] als neue eigenständige Disziplin bzw. als ein neues Forschungsfeld betrieb. Belting ging es vor allem darum, sein eigenes Fach, die Kunstgeschichte, zur führenden Bildwissenschaft unter den seiner Meinung nach diversen Bildwissenschaften [Plural] zu erklären, zu denen er ausweislich der Beiträge des später erschienenen Konferenzbandes auch die Psychoanalyse à la Freud oder Lacan sowie die Literaturwissenschaft rechnete, um nur diese zu nennen. Auf Sachs-Hombach verweist er nur in einer Fußnote (BELTING 2007: 11, Anm. 1 zu einem Interview Sachs-Hombachs von 2004).

So viele Bildwissenschaften also? Ich denke, korrekterweise hätte Belting doch wohl von Disziplinen sprechen müssen, die sich in irgendeiner Form des Bildbegriffs bedienen, manchmal vielleicht sogar Beiträge zum Verständnis des Phänomens ›Bild‹ liefern oder einzelne Bildfragen diskutieren, die das Bild als Bild jedoch auf keinen Fall zum eigentlichen Gegenstand ihres Faches machen. Nach Ansicht von Klaus Sachs-Hombach stand aber genau das auf der epistemischen Agenda. Neben weiteren Pu-

blikationen legten seine schon genannten Sammelbände zur *Bildwissenschaft* den Grundstein zu einer solchen Spezialdisziplin mit der Bildfrage als zentralem Gebiet der Forschungsarbeit (SACHS-HOMBACH 2005a und SACHS-HOMBACH 2005c). Es ging um eine Spezifikation im Fächerspektrum, die es so noch nicht gab und die allein um das Bild kreisen sollte. Der Band *Bildwissenschaft. Disziplinen, Themen, Methoden* betrieb Feldforschung mit Beiträgen zur Frage, was 28 wissenschaftliche Fächer oder Fachgebiete in Deutschland zur Bildfrage beitragen können, darunter auch das Fach ›Allgemeine Rhetorik‹ in Tübingen (KNAPE 2005b). Doch das Kernanliegen war, wie gesagt, die Konstituierung einer eigenständigen ›allgemeinen Bildwissenschaft‹, wie sie in den 1920er-Jahren ansatzweise auch schon von Aby M. Warburg (1866–1929) angedacht worden war (zusammenfassender Überblick zum gesamten Feld bei GÜNZEL/MERSCH 2014).

Bei meinen folgenden Überlegungen geht es nicht um die gerade angesprochene *Bildwissenschaft* mit ihren weit reichenden und vielfältigen Problemen, sondern nur um die *Bildtheorie* im engeren Sinn. So wären etwa Fragen der Bildgeschichte oder der Bildökonomie (Bild als wirtschaftliches Faktum) Gegenstand der Bildwissenschaft, nicht aber der Bildtheorie, die nur das Bild als semiotisches Faktum betrifft. Natürlich gehört die Bildtheorie selbst als Forschungsgebiet ebenfalls in den Zuständigkeitsbereich der Bildwissenschaft. Zu den Archegeten der Bildtheorie zählen Edmund Husserl (1859–1938) und sein Schüler, der Pole Roman Ingarden (1893–1970) mit seiner grundlegenden Abhandlung zum Bild (zu Husserl siehe KNAPE 2020: 237–243; INGARDEN 1962 [1957]).

Kulturalistisch gesehen sind Bilder menschengemacht und besitzen im Sinn der Multimodalitätstheorie eine Schnittstelle ausschließlich zum optisch-visuellen sensorischen Wahrnehmungskanal des Menschen. Wenn ich also am 16. Oktober 2023 im Mittagmagazin von swr-Kultur als Empfehlung für eine Gemäldeausstellung den Satz höre: »Bilder sprechen eine laute Sprache«, dann ist daran theoretisch gesehen so gut wie alles falsch. Bilder sprechen nicht. Und ob es eine Bildsprache gibt, ist noch zu diskutieren. Man kann Bilder auch nicht hören, denn sie sind unakustisch weder laut noch leise. Solche dubios-metaphorisierenden Formulierungen ließen sich als banale vorwissenschaftliche Redeweisen von Journalisten abtun, wenn es diese saloppe und theoretisch fragwürdige Rede nicht auch in der mit Bildern befassten Fachwissenschaft gäbe.

## 2. Probleme der Bildtheorie

David Hilbert, einer der Begründer der modernen Mathematik, wies in seinem berühmten Pariser Vortrag *Mathematische Probleme* aus dem Jahr 1900 darauf hin, dass Forscher immer wieder »eine Musterung der Probleme« ihres Faches und ihres Wissens vornehmen müssen, indem sie sich »die offenen Fragen« ihrer Disziplin vor Augen stellen »und die Probleme überschauen, welche die gegenwärtige Wissenschaft stellt, und deren Lösung wir von der Zukunft erwarten«. Dabei gilt: »Solange ein Wissenszweig Überfluss an Problemen bietet, ist er lebenskräftig; Mangel an Problemen bedeutet Absterben oder Aufhören der selbstständigen Entwicklung. Wie überhaupt jedes menschliche Unternehmen Ziele verfolgt«, so brauche auch die aktuelle Forschung immer wieder »Probleme« als Antriebsfaktor (HILBERT 1900: 253).

Eine »Musterung« von 23 Problemen soll im Folgenden auch für die Bildtheorie unternommen werden. Hilberts methodische Hinweise sind dabei eine Art Leitplanke. Der »Geist« des Theoretikers »schafft aus sich selbst heraus«, schreibt er, »oft ohne erkennbare äußere Anregung allein durch logisches Combiniren, durch Verallgemeinern, Specialisiren, durch Trennen und Sammeln der Begriffe in glücklichster Weise neue und fruchtbare Probleme und tritt dann selbst als der eigentliche Frager in den Vordergrund«. Dabei geht es methodisch um ein »wechselndes Spiel zwischen Denken und Erfahrung«, das unter anderem zur Wahrnehmung »überraschender Analogien« auf »verschiedenen Wissensgebieten« führt. Hilbert ist klar, dass die modernen Wissenschaften immer auch »Importwissenschaften« in dem Sinn sind, dass sie aus Nachbardisziplinen sinnvolle Ergänzungen aufnehmen, dabei theoretische Analogien prüfen, oft durch Erkenntnisimporte ihren eigenen Komplexitätsgrad erhöhen und offene Fragen klären. Durch Analyse der »wirklichen Erscheinungen« der Welt und durch Erschließung von »neuen Wissensgebieten« finden Wissenschaftler »häufig die Antworten auf alte ungelöste Probleme und fördern so zugleich am besten die alten Theorien« (HILBERT 1900: 256f.). Über allem aber steht »die Forderung nach Strenge in der Beweisführung« (HILBERT 1900: 257; »rigor in the proof«; HILBERT 1902: 441) bei der Untersuchung grundlegender »Begriffe« und der »diesen Begriffen zu Grunde liegenden Principien« mit dem Ziel, »ein einfaches und vollständiges System von Axiomen derart festzulegen, daß die Schärfe der neuen Begriffe und ihre Verwendbarkeit zur Deduktion« den alten »Begriffen in keiner Hinsicht nachsteht« (HILBERT 1900: 259).

Für die im Folgenden gebotene Problemübersicht zur Bildtheorie gilt gleichermaßen, was Hilbert für seinen Wissensbereich sagt: »Unermesslich ist die Fülle von Problemen der Mathematik, und sobald ein Problem gelöst ist, tauchen an dessen Stelle zahllose neue Probleme auf. Gestatten Sie mir im Folgenden, gleichsam zur Probe, aus verschiedenen mathematischen Disziplinen einzelne bestimmte Probleme zu nennen, von deren Behandlung eine Förderung der Wissenschaft sich erwarten läßt« (HILBERT 1900: 262).

Zur Bilderfrage hat die in den letzten Jahrzehnten entstandene, reiche wissenschaftliche Literatur sehr viele Fragen aufgerufen und Probleme aufgeworfen. Dabei hat sich eine Reihe theoretischer Probleme als Zentrum der Bilddebatte herausgestellt (vgl. die 15 bildtheoretischen Eckpunkte bei KNAPE 2007: 12–15). Im Folgenden geht es nicht um die Bildwissenschaft als Disziplin mit all ihren Forschungsverzweigungen, sondern ausschließlich um die wissenschaftliche *Bildtheorie*. In den wichtigsten Theoriebereichen besteht auch weiterhin Klärungs- und Diskussionsbedarf. Ich habe sie unter 23 Punkten in meinen Formulierungen für künftige Diskussionen zusammengestellt. Als wissenschaftlich-theoretische Probleme rekurren sie nicht auf alltägliche oder umgangssprachliche Sprachgebräuche oder Bildvorstellungen.

## 2.1 *Erstes Problem: Axiome*

Die Bildtheorie wird im Folgenden als wissenschaftliche Theorie verstanden. »Wenn es sich darum handelt, die Grundlagen einer Wissenschaft zu untersuchen, so hat man ein System von Axiomen aufzustellen«, sagt Hilbert, welches die »elementaren Begriffe jener Wissenschaft« definiert, ihre »Beziehungen« untereinander festlegt und als »Grundlagen« für alle »Aussagen innerhalb des Bereichs der [betreffenden] Wissenschaft« dient (HILBERT 1900: 264). Bislang gibt es solch eine Axiomatik auf dem Gebiet der Bildtheorie noch nicht. Ich schlage Folgendes vor: Als allgemeine axiomatische Prämisse der Bildtheorie hat zu gelten: Es gibt Bilder als Entitäten *sui generis*. Auf das ›Bild‹ als kommunikatives Format beziehen sich weitere spezifische Theorie-Axiome als untergeordnete sinnvolle Prämissen für alle weiteren Spezifikationen und Ableitungen. Es handelt sich dabei um die folgenden fünf weiteren Axiome zum ›Bild‹: *Artifizialität*, *Semiotizität*, *Textlichkeit*, *Sichtbarkeit*, *Bildzeichenstillgestelltheit*. Zu den enthaltenen Implikationen gehören weitere, untergeordnete Prämissen, die an dieser Stelle noch nicht genannt werden.

## 2.2 *Zweites Problem: Artifizialität*

Ein Bild ist ein Artefakt. Mit dem Axiom der Künstlichkeit verbindet sich die Tatsache, dass Bilder gemacht werden und nicht wachsen. Das Bild ist also ein kulturelles, künstlich erzeugtes semiotisches Gebilde: per definitionem nie ein Naturprodukt. Wer also hat uns jeweils das Bild als Zeichengestell hergestellt? Bildermacher und Bilderverwender.

Wichtig für die theoretische Modellierung des Bildes ist seine Zuordnung zu den eigens für die Kommunikation gemachten Typen von Artefakten, die wir Symbole, Zeichen, Texte usw. nennen. Diese Instrumente dienen nicht der biologischen Subsistenzsicherung des Menschen im materiellen Sinn, sondern ausschließlich der Kommunikation. Sie ermöglichen in der symbolischen Interaktion den Austausch von Informationen und konstituieren Virtualität. Was heißt das? Wie alle Texte erzeugen auch Bilder in uns bei ihrer Wahrnehmung rein virtuelle Welten. Das ist der Grundmechanismus der Semiose. Virtualität besteht aus der Simulation von Weltheit als Zeichenrepräsentation in Medien (korporalen oder technischen). Wir imaginieren sie bei der mentalen Bildverarbeitung als *virtual worlds* gemäß semiotischer Zwei-Sphären-Theorie: ›Lebenswelt‹ (Husserl) versus ›Zeichenwelt‹, ›Realwelt‹ versus ›Virtualwelt der Zeichen‹ (zur Virtualitätsfrage siehe KNAPE 1997: 48f. und KNAPE 2019: 43–45; KNAPE 1997: 48f.; zur possible world-theory siehe KRIPKE 1980). All das betrifft nur den ontologischen Status von Zeichenwelten und hat noch nichts mit der Frage von Fiktionalität und Faktizität zu tun.

Umgangssprachlich ist zwar bisweilen auch bei ›natürlich‹ entstandenen Strukturen vom ›Bild‹ die Rede und wir nennen dann etwa den physikalischen Effekt der optischen Reflexion statt ›Spiegelung‹ ein ›Spiegelbild‹, einen Traum ein ›Traumbild‹ oder einen Schatten ein ›Schattenbild‹, doch das ist vortheoretische Rede (vgl. NÖTH 2000: 472f.).

## 2.3 *Drittes Problem: Bildwissen*

Warum können Menschen Bilder verstehen und deuten? Weil die Teilnehmer an Bildkommunikationsprozessen über ein gemeinsames Bildwissen verfügen (Ikonischer Common Ground). Goodman und Elgin nennen es »Bildkompetenz« (GOODMAN/ELGIN 1989: 148–157). Diese ermöglicht es den Menschen bei der Bildinformationsverarbeitung, Bilder, die man vorher

noch nie gesehen hat, auf gewisse Weisen zu verstehen. Bildwissen einer BildCode-Verwendergruppe ist das Äquivalent zum Sprachwissen oder Weltwissen, über das Kommunikationsteilnehmer ebenfalls verfügen. Solche im Körper angelegten Wissensarchive müssen bei der mentalen Verarbeitung von Bildinformationen, bei Produktion und Rezeption von Bildern einbezogen werden.

Unter den Bedingungen der Standard-Frame-Prämisse (siehe Problem 14: Pragmatik) umfasst das Bildwissen (1) Codewissen als Wissen um ein gemeinsam zur Verfügung stehendes Bildzeichensystem, also der Bildzeichenvorrat einer Weltgegend und zugehörige Modelle »syntagmatischer Beziehungen« (zum Begriff GADLER 1992: 152f.), (2) Bildtextwissen und (3) Bildgattungswissen. Die Kenntnis der regionalen BildCodes jenseits weiterer kultureller, nur für die Kommunikation bestimmter SymbolCodes (z. B. Verkehrszeichen) ist fundamental bei der Kommunikation mittels Bild. Was zu welcher Codesorte gehört, muss bisweilen diskutiert werden.

ABBILDUNG 1  
Schlafzimmer



Quelle: Duden. *Bildwörterbuch der deutschen Sprache*. 4. Aufl. Mannheim etc. 1992, 88, Nr. 43 (zugeschnitten)

In der Erziehung speichern Kleinkinder Bildzeichen schon früh beiläufig aufgrund von Zeige- und Benennungshandlungen der Eltern oder mithilfe kindgerechter Bilderbücher. Über den Grundbestand unseres mitteleuropäischen BildCodes könnten sich Nichteuropäer mithilfe des *Bilderdudens* informieren (Abb. 1). Es gibt auch *picturebooks* mit Kodifizierungen von überregional bekannten Bildzeichen (KIEFER 1995; MEADER 1995). Unter Verwendung einer solchen Bildzeichenkodifizierung (eines Bildlexikons) ließen sich dann auf propositionaler Ebene die Interieur-

ABBILDUNG 2

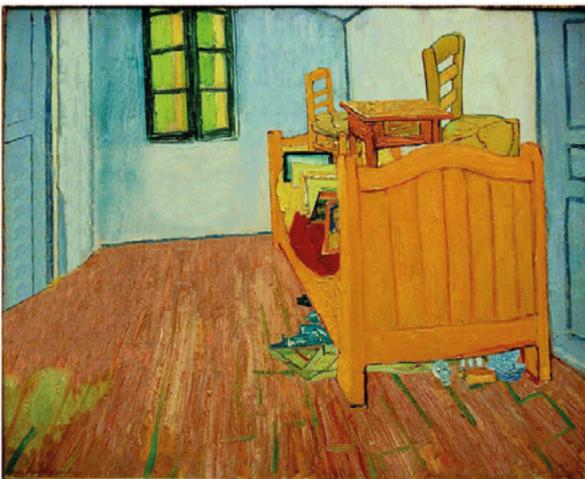
Vincent van Gogh: *Das Schlafzimmer in Arles* (1888)



Quelle: The Yorck Project/Wikimedia: [https://commons.wikimedia.org/wiki/File:Vincent\\_Willem\\_van\\_Gogh\\_137.jpg](https://commons.wikimedia.org/wiki/File:Vincent_Willem_van_Gogh_137.jpg)

ABBILDUNG 3

Ursus Wehrli: *Kunst aufräumen*



Quelle: Wehrli 2002: 21

bestandteile von Vincent van Goghs *Schlafzimmer* als semantische Elemente einer Zeichenkomposition, realisiert als Bildzeichengestell, identifizieren (Abb. 2): Stuhl, Handtuch, Tisch, Spiegel, Bett. Nichts bewegt sich in Van Goghs *still picture*, so wie in jedem anderen Bild auch. Wenn sich jedoch die räumliche Anordnung der Bildzeichen ändert (Abb. 3), dann ändert sich auch die Semantik des Bildes in toto.

#### 2.4 *Viertes Problem: Bindestrich-Theorien*

Die Bildtheorie ist wissenschaftlich und eigenständig. Sie ist Teil der Bildwissenschaft. Zugleich steht sie in einem Nachbarschaftsverhältnis zu Theorie- und Wissenschaftsfeldern, die sich ebenfalls mit Bildfragen beschäftigen, dabei jedoch ihre eigenen Perspektiven auf die Bilderfrage entwickeln. Für sie werden Bindestrich-Begriffe wie ›Bild-Philosophie‹, ›Bild-Psychologie‹, ›Bild-Recht‹, ›Bild-Kunst‹ oder ›Bild-Rhetorik‹ etc. verwendet. Diese an die Bildwissenschaft angrenzenden Gebiete sind wichtig, weil Importe aus ihnen das Theoriegebäude der Bildtheorie anreichern können.

#### 2.5 *Fünftes Problem: Definition*

Was ist ein Bild? Diese Frage wurde und wird in der *scientific community* kontinuierlich diskutiert. Der wissenschaftliche Forschungsgegenstand ›Bild‹ braucht eine wissenschaftliche Theorie. Vor diesem Hintergrund hat Klaus Sachs-Hombach bereits im Jahr 2005 den Ansatz eines Definitionsumrisses geliefert, der alle weiteren wissenschaftlichen Überlegungen in eine sinnvolle Richtung hätte führen können und müssen, was aber keineswegs geschah. »Bilder im engen Sinn« sind aufzufassen, so Sachs Hombach, als »Gegenstände, die materiell, in der Regel visuell wahrnehmbar, artifizuell und relativ dauerhaft sind« (SACHS-HOMBACH 2005b: 12f.). Diese Bestimmung war zu ihrer Zeit sinnvoll und wissenschaftlich geboten. Dennoch haben viele Benutzer des Bildbegriffs seit 2005 dieses ausbaufähige Vorverständnis nicht wahrgenommen, geschweige denn weiterentwickelt, und stattdessen an ihm vorbeigedacht; keineswegs zum Nutzen der Sache. Inzwischen hat sich die Diskussion weiterentwickelt und man kann heute in aller Kürze sagen: Ein Bild ist ein Text aus stillstehenden Bildzeichen in kommunikativer Absicht.

Aus dieser Definition ergeben sich eine Reihe von Merkmalen, die ein Bild ausmachen (KNAPE 2016). 1. Zeichenkomplex (Text): Als Text vermittelt ein Bild Bedeutungen. Das ist seine Zweckbestimmung. 2. Bildzeichen: Die Bildzeichen sind die entscheidenden bedeutungsgenerierenden Elemente des Bildes. 3. Ordnung: Die Zeichen sind im Bildraum nach semantischen Erfordernissen zu höher organisierten Strukturen geordnet, damit Bedeutungen komplexerer Art kommuniziert werden können.<sup>1</sup> Die Bildzeichen stehen nach semantischen Kriterien im Bildraum und im Verbund zusammengestellt als Gestell still. 4. Begrenzung: Innertextliche, pragmatische oder mediale Strukturen begrenzen ein Bild als Einzeltext und unterscheiden es dadurch von einem Hyper- oder Superbildtext (zu den verschiedenen Rahmungsdimensionen siehe SCHAPIRO 1995). Bilderfolgen oder Bildersequenzen und natürlich auch Filme sind keine ›Bilder‹. 5. Kommunikatives Instrument: Bilder sind Bedeutungsträger und dienen damit kommunikativen Zwecken. Bilder sind keine Dialoge. Das Bild gehorcht dem interaktional-kommunikativen Basismodus des Monologs (wie schon Platon im 4. Jh. v.u.Z. bemerkte, *Phaidros* 275d) mit unidirektionalem, unilateralem und uniintentionalem Anspruch an angesprochene (besser: ›angebilderte‹) Adressaten.

Nicht bei allen, aber bei vielen Bildgattungen spielt auch die Farbgebung eine wichtige semiotische (!) bzw. künstlerische Rolle. Für eine allgemeingültige (!) Bilddefinition ist die Farbe jedoch als Kriterium auszuschließen, weil es viele Bildgattungen gibt, die nur mit Schwarz-Weiß-Kontrasten arbeiten und auf die folglich ein Definiens ›Farbe‹ nicht anwendbar wäre.

## 2.6 *Sechstes Problem: Funktionstypen, Format- und Gattungsfragen*

Das Bild ist unter den Visualisierungsformaten das prominenteste Format (KNAPE 2023b: 20f.). Die gesamte Bildgattungsfrage ist zwar noch im Fluss, doch kann man grundsätzlich zwei Funktionstypen unterscheiden: *Fiktivbilder* arbeiten auf der Basis von typisierten Gestaltformen, die jede Art von Bildphantasie erlauben, *Abbilder* hingegen wollen als Faktualbilder die

1 Das darf nicht mit dem unbegrifflichen, ästhetischen Zugang zu Kunstwerken verwechselt werden (KNAPE 2023c: 42–44 und 85f.); vgl. dagegen BOEHM 1994: 21f.

konkrete Wiedererkennung von außersemiotischen Realitäten auf Basis derselben Gestaltformen gewährleisten (z. B. bei imitativen Porträts). In der konkreten Kommunikation legt der Frame die jeweilige Funktion fest (Abbild oder nicht). Bilder sind aus Bildzeichen komponierte Zeichenkomplexe, die beide Kommunikationsfunktionen ermöglichen. Die Frage, wie sich der Realitätsgehalt bei datengebundenen Abbildern stärker gegenüber dem Fiktionsgehalt bei Autonombildern herausarbeiten lässt, wird die Bildtheorie als Problem in Zukunft weiter beschäftigen.

Zu den immer wieder auftauchenden Unverständnissen alltäglicher Bildkonzepte gehört die Annahme, Bilder seien schon immer und in erster Linie »Abbilder«. Das aber ist ein Irrtum. Es gibt (fiktive) Bilder, die nichts »abbilden«, dennoch aber eine Semantik haben. Bildtexturen entstehen in uns als Verarbeitungsreflexe auf äußere Zeichenangebote und unter Einsatz unseres inneren Bildwissens als Zeichenwissen, so wie lautsprachliche Texturen unter Einsatz unseres Sprachwissens entstehen. Aufgrund bestimmter Einbettung in kommunikative Zusammenhänge können Bilder Vielfältiges leisten. Für sich genommen sind sie freilich regelmäßig informationell unterspezifiziert aufgrund der Eigenart der Bildzeichen. Man kann hier von der »Unschärferelation des Ikonischen« sprechen (KNAPE 2025b: 362).

Fundamental ist die eben schon genannte Unterscheidung in zwei Funktionaltypen, denn es gibt Bilder, die etwas abbilden, und es gibt Bilder, die nichts abbilden, dennoch aber Bedeutung haben (KNAPE 2019: 52f.; 2020; 2023a und 2025a: 122):

1. Autonombild oder Fiktionsbild als bloßes *Verisimile*, als nur Wahrähnliches ohne bildexterne Objektreferenz. Das verisimilische Bild ist ein für sich stehender Bildtext ohne textexternen Referenten. Diesem Funktionstyp des Bildes ordnen wir bei der mentalen Zeichenverarbeitung den Status eines künstlichen Phantasmas zu (Fiktionalität). Die Bildlinge der Verisimiles sind demzufolge virtuell existent (wie alles auf der theoretisch postulierten Zeichenebene Angesiedelte), darüber hinaus aber auch noch *fiktiv* (erfunden). Wir können hier von Objekt-Präsenz-Illusion ohne Deixis auf die *actual world* und vom Bild als reiner Fiktion sprechen (folglich ohne Dokumentarcharakter). Damit eröffnet das Autonombild das Reich der Bildfreiheit, weil nur hier unter Rückgriff auf das erlernte Bildwissen neue fiktionale Bildwelten, d. h. außenreferenzlose *possible worlds* erschaffen werden können. Dies ist möglich, weil wir Menschen über Bildwissen in Gestalt eines BildCodes verfügen, der angesichts fiktionaler Bildtextangebote

im Denken aktualisiert wird. Autonombilder sind »im Unterschied zum bloßen Abbild ein Erzeugnis des Geistes und nicht ein Niederschlag der physischen Wirklichkeit«; sie sind ein autonomes Zeigen »der Struktur der Dinge«, wie sie ein Bildermacher intendiert hat (WEDEWER 1985: 183).

2. Abbild oder Faktenbild als *Simulacrum*, als Gleichbild mit hohem Ähnlichkeitswert hinsichtlich einer äußeren Objektreferenz der *actual world*, also mit Verweis auf etwas physikalisch konkret Sichtbares, das sich außerhalb des Bildes befindet (Realitätsimitation und Faktizität). Ein Bild kann die hier in Frage stehende kommunikative Funktion eines *Simulacrum*s bekommen (etwa als Porträt), (»*Simulacrum*« als »Abbild« schon 1764 bei Lambert; siehe LAMBERT 1965 [1764]: Tom. II/1, 9; KNAPE 2025a) wenn die Textur so gestaltet ist, dass sie in der Summe hochgradig merkmalsähnlich zu realen, außertextlichen Sachverhalten ist. Der realitätsmotivierte *token* regiert, der *type* tritt zurück. *Simulacra* fingieren nicht, haben aber dennoch als Teil der Zeichenrealität Virtualitätscharakter. Demzufolge sind die identifizierbaren Dinge des simulacrischen Bildes im Wahrnehmungszusammenhang nur virtuell existent. Jedes Abbild von Objekten erfüllt damit die Bedingungen der »Zweiteit« einer Ding-»Inszenierung« (KNAPE 2017; 2025a). Erst bestimmte Gattungsmarkierungen oder Frame-Ansagen definieren den kommunikativen Status eines Bildes als Gleichbild, »*Simulacrum*« oder »Abbild«. Damit wird ein Bild hinsichtlich seiner Funktion als Dokument ausgewiesen (zu Dokumentarfunktion SUSANKA 2015: 86f.; KNAPE 2020).

Ansonsten ist das Textsorten- oder Gattungssystem der Bilder in unterschiedlichen Disziplinen divers ausdifferenziert. Die Kunsthistoriker etwa teilen Bildgattungen nach Themen ein: Landschaftsbild, Seestück, Schlachtenbild, Porträt, Stilleben usw. Juristen sortieren nach dem Dokumentenstatus: Passbild, Steckbriefbild, Verkehrskamera- oder Video-Beweisfoto usw. Insgesamt ist die Theorie der Bildgattungen und Bildtypen wissenschaftlich gesehen noch nicht weit ausgearbeitet (SCHMITZ 2007: 421).

## 2.7 Siebtes Problem: Grammatik

Wenn in Kommunikationen mit Hilfe von Bildern Verständigung erfolgen soll, müssen Bildproduzenten- und Rezipienten über denselben BildCode verfügen (KNAPE 2013a: 424–430). Schließt dieser aber auch so etwas wie Bildgrammatiken ein (NÖTH 2000: 478)? Was ist eine Grammatik? Eine

Grammatik ist im strengeren Sinn dieses Terminus technicus ein Orthosystem, also das Richtigkeits-Regelwerk einer Sprache. Jeder, der schon einmal eine Fremdsprache gelernt hat, weiß, was die grammatische Richtig/Falsch-Opposition bedeutet. Maßstab ist dabei immer eine festgelegte Zeichengebrauchs-Konvention als Kollektivgebrauchsnorm unter den Bedingungen der Standard-Frame-Prämisse (siehe Problem 14: Pragmatik; NÖTH 2000: 341). Intersemiotisch gesehen gibt es allerdings keinen einheitlichen Typ von Regelwerken für alle Zeichensysteme. Linguistisch gesehen ist eine Grammatik ein Regelwerk immer nur für eine Einzelsprache, das die Elemente dieser Sprache definiert und die verbindlichen Regeln ihrer Verwendung bei der Konstruktion sprachrichtiger Strukturen nicht auf Textebene, sondern auf der Ebene von ›Textbausteinen‹ (Wörtern, Sätzen) angibt (KNAPE 2024: 94–101). Solch eine Textbaustein-Grammatik bis zur Ebene des ›Satzes‹ kann es beim Bild nicht geben (zur kontroversen Diskussion dieser Frage siehe NÖTH 2000: 479f.; vgl. WEDEWER 1985: 187–200).

Wichtig ist, wie gesagt, für das Verständnis von Grammatikalität, dass Grammatiken Orthosysteme, also Richtigkeitssysteme nach Korrektheitskriterien sind. Das theoretische Hauptproblem beim Zusammenhang von Bild und Grammatikalität liegt, dies muss noch einmal betont werden, in der Tatsache begründet, dass Grammatiken aufs Sprachsystem als ›Baustein- und Regellieferant‹ (wie man sagen könnte), nicht auf die Ebene des Textes als Ebene der ›Textgebäudekonstruktionen‹ mit übersummativen Bedeutungen bezogen sind (KNAPE 2024: 94–101, 105–114). Bei den Lautsprachen gelten Grammatiken normalerweise nur auf der semantischen Baustein-Ebene unterhalb der Textebene (sprachliche Minimalelemente, Morphem/Wort, Wortgruppe, ›Klausel‹, Satz sowie Lautsystem; COSERIU 1980: 30).

Ob es so etwas wie ›Textgrammatiken‹ überhaupt geben kann, ist umstritten, weil der Text eigentlich ein nur pragmatisch (also von kommunikativen Zielen) bestimmter und okkasionell auftretender Freiraum ohne systemischen Charakter ist.<sup>2</sup> Trotz aller Bedenken wird in der Forschungs-

2 Eugenio Coseriu, einer der Begründer der sprachwissenschaftlichen Textlinguistik, unternahm 1980 dennoch den Versuch einer »Rechtfertigung der ›transphrastischen Grammatik‹ für lautsprachliche Texte. Damit meint er Regeln, die satzüberschreitend gelten sollen. Allerdings bleibt unklar und zweifelhaft, welchen theoretischen Status diese Regeln haben, ob sie mit dem Begriff ›Grammatik‹ passend bezeichnet sind und ob sie sich tatsächlich auch nur spezifisch auf die Textebene und Textorganisation beziehen (COSERIU 1980: 16–29). Die von Coseriu genannten satzübergreifenden Regeln betreffen nämlich idiomatische Formeln (einschließlich ›unmöglicher Sätze‹), rhetorische Figuren, Redemodalitäten und regulierte

literatur, ausgehend von einem vorwissenschaftlichen Grammatikkonzept, die Existenz einer ›Bildgrammatik‹ postuliert. Für die Bildtheorie haben Kress und van Leeuwen 1996 den Versuch unternommen, eine *Grammar of Visual Design* zu entwickeln. Dabei geht es nicht nur um ›Design‹ im engeren Sinn (KNAPE 2019: 99–147), sondern auch ums ›Lesen von Bildern‹ (*Reading Images*), was schon im Titel darauf hindeutet, dass die Autoren kein Verständnis für die Analog-digital-Differenzen haben und meinen, das schriftbezogene ›Lesen‹ auch aufs Bild übertragen zu können. Nicht viel besser steht es dann um den verwendeten Grammatikbegriff. Die Autoren distanzieren sich von der Orthokomponente des üblichen Grammatikverständnisses und räumen ein, dass es in ihrem Buch nicht wirklich um Grammatikalität im terminologischen Sinn (mit Korrektheitsprämissen) geht, sondern um die Zusammenstellung von herrschenden Gewohnheiten oder Best Practices als optimale Vorgehensweisen in Design und Bildgestaltung (KRESS/VAN LEEUWEN 1996: 1–3). Warum dann aber die missbräuchliche Verwendung des Grammatikbegriffs für Bildhaftes, fragt sich der Theoretiker? Offenkundig wollen sich Kress und van Leeuwen damit jenen zu erkennen geben, denen Kategoriefehler ebenso egal sind wie auch ihnen und die ebenfalls keinen Wert auf theoretisch strenges Begriffsverständnis legen. Nach dieser Sicht der Dinge wäre alles Grammatik, das irgendwie Ordnungen einfordert. Das allerdings ist theoretisch fragwürdig. Dennoch hat sich diese Sicht gehalten. So behauptet etwa Descola noch 2021, die Lütticher Semiotiker der Gruppe  $\mu$  (Groupe  $\mu$ ) hätten 1992 in ihrem *Traité du signe visuel* (GROUPE  $\mu$  1992) eine »allgemeine Grammatik des Bildes« ausgearbeitet. Tatsächlich handelt es sich bei der Theorie des *Traité* jedoch um eine abstrakte, bildbezogene strukturalistische Figurenlehre (aus dem Geist der rhetorischen Figurenlehre) auf der Basis einer komplizierten Zeichentheorie des Bildzeichens, über die Descola dann – widersprüchlich genug – selbst sagt, dass sie sich nicht etwa von anderen Grammatiken, sondern von den »üblichen semiotischen [!] Theorien« abhebe (DESCOLA 2023: 657).

Insgesamt gilt für die Grammatikfrage, wie auch für alle anderen Bereiche der Bildtheorie, dass es zukünftiger Forschung vorbehalten bleibt, die theoretischen Modellierungen auf den Prüfstein zu stellen und wei-

Funktionsstellen im Text. Auch Ausdrucksweisen, die sowohl im Satz gelten, als auch »über den Satz hinausweisen (Partikeln, Topikalisierung, Ersetzung)«, die also nicht textspezifisch sind, bezieht er ein (COSERIU 1980: 28).

terzuentwickeln. Das betrifft hier insbesondere die Frage, wie die Richtigkeitsbedingungen in den Bildgestalten theoretisch noch differenzierter modelliert werden können.

## 2.8 *Achtes Problem: Informationsleistung, spezifisch*

Bilder zeigen Zeichen-Gestelle, können nur Zusammenstellungen von Bildzeichen darstellen. Ihre informationelle Leistungsfähigkeit ist gegenüber Texten aus lautsprachlichen Elementen sehr eingeschränkt. So haben sie als *stills* etwa keine oder nur eingeschränkte Ausdrucksmöglichkeiten für Kausalitäten, Handlungen oder Bewegungen. Betrachter brauchen für die Interpretation also viele Zusatzinformationen, die über die Frames eingeholt werden müssen (siehe Problem 14: Pragmatik; ein Beispiel bei BÄTSCHMANN 1989). Allerdings ist der Bildfunktionstyp *Simulacrum* mit dem ganz besonderen Leistungspotenzial der Gestalt-Realitätsimitation ausgestattet, was lautsprachliche Texte nicht vermögen (siehe Problem 6: Funktionstypen, Format- und Gattungsfragen).

De Beaugrande und Dressler verstehen 1981 unter Informativität eines Textes »das Ausmaß der Erwartetheit bzw. Unerwartetheit oder Bekanntheit bzw. Unbekanntheit/Ungewißheit der Inhalte der dargebotenen Textelemente« (DE BEAUGRANDE/DRESSLER 1981: Kap. I.17–I.18; HÖRMANN 1978: 27). Hintergrund ist das allgemeine Informativitätsprinzip des Textes. Damit ist letztlich, so De Beaugrande und Dressler, jenes Ausmaß gemeint, »bis zu dem eine Darbietung für den Rezipienten neu oder unerwartet ist« und damit »Aufmerksamkeit« erzeugt (DE BEAUGRANDE/DRESSLER 1981: Kap. VII.1). Unerwartet ist eine Information, wenn sie von geringer »kontextueller Wahrscheinlichkeit« im kommunikativen Einbettungsrahmen ist, d. h. wenn sie zu jenen Formulierungsoptionen (Sprachelementen, Äußerungsarten und -muster, Inhalten usw.) gehört, die in Standard-Kontexten selten vorkommen.

Bilder, die per definitionem aus Bildzeichengestalten bestehen, können nur über Visuelles, Gestalthaftes, Lichthaftes (hell/dunkel) und Stillgestelltes informieren.<sup>3</sup> Für die Abstrakta, Kollektiva, Verben (WEDEWER

3 Vgl. NÖTH 2000: 482. Platon abstrahiert in dieser Frage im 4. Jh. v.u.Z. ontologisch radikal und verweist in der *Politeia* auf das prinzipiell ontologische Defizitäre des Bildermachens in

1985: 135–157) oder gar die unsemantischen (rein satzfunktionalen) Kopula der Lautsprachen gibt es auf Zeichenebene im Bild keine Äquivalente. Erst im Syntagma entstehen übersummativ Semantiken im Bild. Viele Bildgattungen können zudem Informationen über Farbverhältnisse liefern.

Die stillgestellten Zeichenkonstellationen des Bildes sind räumlich, keinesfalls aber zeitlich organisiert. Sollen Bilder dennoch etwas über zeitliche, nicht-visuelle oder nicht-konkrete Verhältnisse mitteilen, vielleicht sogar über Abstraktionen oder Ideen, müssen sie andere kulturelle SymbolCodes oder die Schrift als Notationsform von Sprache einbeziehen, um Zusatzinformationen zu erzeugen; es entsteht ein semiotisches Hybridbild als Visualisierungsformat per CodeMix (etwa im Fall von Funktionstypen wie Simulacra oder Bildallegorien mit erklärendem Schriftband; KNAPE 2023a, b). In solchen Fällen lässt sich Nicht-Visuell-Abstraktes im Bild oft auch mithilfe komplexer Bildsyntagmen darstellen. In jedem Fall müssen dann aber Frame-Prämissen und Einbettungsinformationen der bildlichen Unterspezifikation bei der Interpretation zu Hilfe kommen.

## 2.9 Neuntes Problem: Kommunikabilität

»Ein Bild ist ein kommunikatives Faktum« (KNAPE 2007: 12). Dieses bildtheoretische Subaxiom ergibt sich aus dem Axiom der Semiotizität. Demnach gilt für Bilder die Bedingung, immer im Dienst der Kommunikation zu stehen, d. h. der symbolischen Interaktion als Adpragmatisierungs-Frame (griech. *pragma* = Handlung; zum Handlungsbegriff siehe JÄGER 1981: 33–47). Der Begriff ›Bild‹ gilt nicht für Naturphänomene, nicht für zufällig entstandene Gebilde oder für solche ohne kommunikative Einbettung. Wenn religiöse ›Kultbilder‹ verehrt werden, geht es im Kern nicht um das Bild als Bild im theoretischen Sinn, sondern um die Tatsache des ›Kultobjekt‹-Charakters eines Mediums als Bildträger (Bildtafel, Gemälde

Hinblick auf Nachahmung, wenn er betont, dass ein Bild immer nur das »Erscheinende« (*phainómenon*), nie das »Seiende« (*ón*) in seiner Fülle zeigen kann: »Gar weit also von der Wahrheit ist die simulierende Nachbilderei (*mimēsis*); und deshalb, wie es scheint, macht sie auch alles [ohne jede Zurückhaltung], weil sie von jedem nur ein Weniges trifft und das im Schattenbild (*eidólon*)« (598 b–c; Übers. n. Schleiermacher). Platon hat damit grundsätzliche Zweifel am Seinscharakter der Phantasmen eines Bildes angemeldet und einen Gegensatz zwischen der realen Existenz eines Dings außerhalb eines Bildes und einem im Bild erkennbaren Ding konstatiert (vgl. DELEUZE 1993: 259–267 und 311–324).

usw.; DANTO 1991: 242f.). Der Kommunikationsmodus der Bildkommunikation ist die Deixis stillgestellter Bildzeichenzusammenstellungen.

### 2.10 *Zehntes Problem: Kunstfrage*

Bildkünstler haben das menschliche Wissen um die Darstellungsmöglichkeiten des Bildes durch ihre Experimente und Ausdrucksversuche entscheidend vorangetrieben. Zukünftige Forschung wird die historischen Stufen dieser seit Urzeiten bestehenden Entwicklung weiterhin auch diachron, aber mit strukturalistischem Blick erforschen müssen.

Die synchrone Bildtheorie geht zunächst von den Bedingungen der Standard-Frame-Prämisse aus (siehe Probleme 7: Grammatik und 14: Pragmatik). Sowie ein Fall von Lizenzkommunikation vorliegt (etwa von Kunst die Rede ist), kommen andere Regulative ins Spiel. Dementsprechend ist nicht jedes Gemälde ein ›Bild‹, (vgl. WEDEWER 1985: 206) nicht jedes Bild (etwa eine Urlaubsfotografie) ›Kunst‹ und die Bildwissenschaft ist keine Kunstwissenschaft. Dennoch ist immer wieder von unverständiger Seite zu hören, die wissenschaftliche Bildtheorie solle sich auch der Kunstfrage zuwenden (also zur Kunsttheorie werden). Das aber wäre etwa so – um den Vergleich zu einer Paralleldisziplin der Bildwissenschaft heranzuziehen – als ob man von der Linguistik forderte, sie solle sich nicht mit der Sprache als Sprache befassen, sondern ihr Feld verlassen und sich literaturwissenschaftlichen und ästhetischen Fragen zuwenden. Ästhetische Fragen kommen bildtheoretisch gesehen nur ins Spiel, wenn sich der Standardframe ändert und damit die Pragmatik.

### 2.11 *Elfte Problem: Medialitätsmodus*

Medien materialisieren Texte, also auch Bilder (KNAPE 2010). Medien sind als Zeichenträger im terminologisch strengeren Sinn (KNAPE 2013b: 251–269) die materielle Ermöglichungsbedingung von Text und seiner Körperexteriorisierung, d. h. technische Medien ermöglichen es, dass Text außerhalb menschlicher Körper manifest wird und in die sozialen Interaktionen eingespeist werden kann. Generell sind Medien »Tragflächen von verbalen oder nonverbalen Texten« (KNAPE 2000a: 62); somit sind Medien die Vorbedingung der Bilder (HUBER 2004: 71); »Medien als Körper der Bilder«

(SCHULZ 2005: 122–124). Demgegenüber ist ein Bild ein Text und damit theoretisch gesehen ein rein semiotisches Phänomen, egal, wie gut oder schlecht es aufgrund der Medialisierungsqualität wahrgenommen werden kann. Texte sind nach semantischen Kriterien und pragmatischen Kommunikationsbedürfnissen organisiert.

Zu den theoretischen Abstraktionen der Bildtheorie gehört also, dass Bildtheorie keine Medientheorie ist. Viele Wahrnehmungseffekte (!) sind beim Bilderleben allerdings rein medial bedingt, also nicht-semiotisch. Dazu gehören alle Fragen der Wahrnehmungsstörung, »visuelles Rauschen« oder »Unschärfe«. Begriffe wie Leinwand, Grundierung, »Farbkontrast« oder »Farbflecken« (BÜHLER 1934: 165), Verwischung, Schlieren, Pixel (NAKE 2005) oder Rasterung usw. sind auf der theoretischen Ebene des Mediums abzuhandeln (KNAPE 2010: 86–91).

Ein zentraler medialer Faktor ist die bildformat-typische *Zweidimensionalität* der Zeichenmedialisierung visueller Informationen durch den Bildträger, den wir im Deutschen aufgrund unseres alltäglich eingeschränkten Wortschatzes ebenfalls oft einfach nur »Bild« (engl. *picture*) nennen. Technisch kann die zweidimensionale Flächigkeit auf unterschiedlichste Art und Weise bedingt sein. Auf der analytisch davon zu trennenden Textebene sind im eigentlichen semiotischen Bild (engl. *image*) natürlich dreidimensionale Illusionskonstruktionen möglich (Perspektive usw.). Neuronal im Menschen verankerte Imaginationen fallen per definitionem nicht ins Gebiet der Bildtheorie, stehen im Medialisierungszusammenhang auf einem ganz anderen Blatt und gehören systematisch gesehen als physiologisches Black-box-Problem zur Abteilung »Bildverarbeitung« (vgl. GOTTSCHLING 2005).

Als materielle Ermöglichungsbedingungen von Bildern kommen alle denkbaren Medien in Betracht, von einer Hauswand über Papier bis hin zum Bildschirm eines Computers (KNAPE 2005a, 2010). Die ständig neu auftretenden Möglichkeiten technischer Bildgebungsverfahren gehören zum Medienkomplex (Stichwort: Vilém Flussers »Technobilder«; KNAPE 2000b). Technische Neuerungen verändern zwar nicht die Abstraktionen der Basistheorie des Bildes, doch ermöglichen sie neue Funktionstypen, Gattungen und Informationsleistungen des Bildes sowie generell neue Visualisierungsformen und -formate. Digitale Animationstechniken haben neue Bildrealisierungsmöglichkeiten geschaffen.

Bei der älteren Medientechnik gehört zu den medial-technischen Bedingungen der Erzeugung von Bildzeichen, dass sich im Moment der Produktion vor den Aufnahmegegeräten physikalisch sichtbare, dinglich-reale

Referenten befinden (Menschen, Dinge, Räume). Im Bild werden sie dann zu Bildzeichen. Die damit nur indirekt verbundene Frage der Fiktion und eines sonderkommunikativen Arrangements muss in diesen Fällen der Frame klären.

## 2.12 Zwölftes Problem: Narrativitätsfrage

Der Kern eines Narrativs liegt informationell vor, »wenn in einem Text als Information ein erkennbarer Zusammenhang zwischen Akteur + Aktion + Aktionszeit auftritt, und sei es auch nur aufgrund der Darbietung des Textes, ohne dass ein Kausalitätsindikator zwingend gefordert wäre« (KNAPE 2025b: 370f.; mit weiteren fundamentalen icono-narratologischen Theoriebausteinen). Auf Realisierungsebene zentral ist hier die Differenz von

1. *Erzählen* (engl. *narration*) im Sinne einer pragmatisch-kommunikativen Handlung mittels medialer Performanz als Medialisierungsphänomen sowie
2. *Erzählung* (engl. *narrative*) als erkennbare textliche Superstruktur, mit allen daran hängenden Implikationen (KNAPE 2025b: 345).

Ad (1): Bilder können nicht im Sinn von Performanz ›erzählen‹. Die 1978 veröffentlichte transmediale Erzähltheorie von Seymour Chatman bietet hier erste Hinweise, legt für uns aber auch das Problem offen (CHATMAN 1978). Er unterscheidet *spatial media* und *temporal media* und meint damit das, was in unserem theoretischen Zusammenhang *Raumtexturen* und *Zeittexturen* heißt. Was ist damit gemeint? Bilder und Skulpturen haben als *stills* selbst keine texturale Zeitebene (siehe Problem 17: Still-Theorem). Natürlich benötigt der Betrachter Zeit beim Betrachten eines Bildes, doch da befinden wir uns auf der analytischen Wahrnehmungsebene eines Akteurs, nicht auf der Bildebene. Wenn wir uns auf die Ebene des Kommunikationsinstruments *Bild* begeben, stellen wir fest, dass es dort keine Zeitdimension gibt, kein »temporal ›program‹ inscribed in the work«, wie Chatman zurecht sagt (CHATMAN 1990: 3). Die fließenden, zeitgebundenen Texturen hingegen erstrecken sich sequenziell und zugleich performativ in der Zeit. Das gilt für Texte menschlich-lautsprachlicher Rede genauso wie für den Film. Die doppelte Zeitachse des Narrativen (Erzählzeit und erzählte Zeit), das wird nun sofort klar, führt zu theoretischen Schwierigkeiten bei einem Bild, das per definitionem ein *still* ist.

Ad (2): Die Existenz des bildlichen *narrative* (Erzählung) ist umstritten (VEITS 2020: 124f.). Den wichtigsten Vorschlag zur Lösung dieser schwierigen Frage hat Marianne Wunsch 1999 gemacht (WÜNSCH 1999). Ihr Ansatz ist überzeugend und fügt sich gut zu den kognitiven Verarbeitungsansätzen der ›narrativen Psychologie‹ (VEITS 2020: 126f.). Wunsch geht zunächst völlig zu Recht von der strengen semiotischen Voraussetzung aus, dass eine kommunikative Äußerung (z. B. ein verbaler Text oder eine Filmtextrur) nur dann eine narrative Struktur hat, wenn bei ihr die tragenden Narrativitätskomponenten (Akteur und Aktion auf der Aktions-Zeitachse) als semantisch relevante Komponenten gegeben sind: (1) »zwei Zustände derselben menschlichen oder nicht-menschlichen Entität« und (2) »eine zustandsverändernde Operation, eine Transformation also«, die sich als Bedeutung des Textes erkennen lassen (WÜNSCH 1999: 328). Genau das aber kommt bei einem Bild als *still* nicht vor, dessen Zeichen weder linear noch zeitsequenziell organisiert sind, weswegen man im Bild auch keine Zeit messen kann. Bei der Lösung des Problems geht Wunsch zurecht davon aus, dass Bilder nach bestimmten Regeln geordnete Zeichenkomplexe sind, die eine Semantik höherer Ordnung aufbauen, wie wir sagen. Unter ihnen gibt es einige, die mittels des erschließbaren *plots* Hinweise auf eine »temporale Komponente« als semantisches Kommunikat geben, d. h. sie müssen »Merkmale derart haben, dass wir aufgrund unseres kulturellen Wissens zu folgern vermögen, der im Bild repräsentierte Zustand« der gezeigten Objekte im Bild »sei früher ein anderer gewesen oder werde später ein anderer sein« können (WÜNSCH 1999: 331). Das könnte etwa durch kodierte Posen oder Gesten von menschlichen Figuren ausgedrückt sein (KNAPE 2013a: 434–442). Diese Bedingung ist notwendig, aber noch nicht hinreichend. Es geht nicht nur um die Temporalfrage, sondern auch um die Handlungsfrage als solche. Das Bild kann eine Handlung »nicht *darstellen*, sondern nur *indizieren*, indem es aus den sukzessiven Teilschritten und Zeitpunkten der Handlung einen charakteristischen Moment auswählt, aufgrund dessen wir erschließen können und müssen, was diesem Zeitpunkt vorangegangen ist und/oder was ihm folgen wird« (WÜNSCH 1999: 334). Wir haben es also auf der einen Seite mit der Textur zu tun, die uns im Fall des Bildes nur etwas unbewegt zeigt (etwa menschliche Posen), und auf der anderen Seite beim Adressaten mit einem Perzeptionsakt und dem Prozess kognitiver Verarbeitung der Informationen unter Einbeziehung von rahmendem Bild- und Weltwissen (Zusammenhang von *images mentales* und *images objets* bei BELTING 1998: 35; siehe auch KNAPE 2025b: 361).

Das entscheidende Verfahren zur Ermittlung narrativer Verhältnisse im Bild ist der »Ekphrasistest«, bei dem durch Transnotation die vorhandenen »Erzählungs«-Strukturen offen gelegt werden können. Allerdings ist dabei auf die mit der Differenz von »Elementarekphrasis« (Zustandsbeschreibung) und »Projektionsekphrasis« (Handlungsvermutung) verbundenen methodischen Probleme zu achten (KNAPE 2025b: 359–365; siehe auch Problem 21: Transnotation).

### 2.13 Dreizehntes Problem: Notation

Visuelle Informationen können nur über die visuellen Formate vermittelt werden, etwa durch das »Bild« (KNAPE 2023b: 20f.). Die Vertreter der *Actor-Network-Theory* Bruno Latour und Steve Woolgar haben die Ansicht radikalisiert, dass nur auf dem Wege materieller Medialisierung eine Information existent ist, und sei es auch nur als neuronal inskribiertes Wissensselement.<sup>4</sup> Nach der *inscription theory* ist auch die Wissenschaft ein System zur Produktion von Inskriptionen und nicht von rein geistig existierenden Theorien (LATOUR/WOOLGAR 1979; FERRARIS 2007). Nur die Textnotate spielen als konkrete Entitäten eine konkrete Rolle im wissenschaftlichen Interaktionsprozess, nicht etwa abstrakte Entitäten wie wissenschaftliche Theorien, die ihre Existenz vermeintlich im luftleeren Raum oder noch deutlicher: als rein geistige Formationen führen. Dieser Ansatz, alles Informationelle in der Welt mit seiner materiellen Existenzform, mit der für seine Formulierung je konkret gegebenen medialen Entität gleichzusetzen, ist eine Weiterentwicklung von Derridas »Grammatologie«-Thesen. Demnach gilt der Grundsatz: *Nihil extra medium* (KNAPE 2013b: 253), dem man prinzipiell zustimmen kann. Doch ist diese Einsicht eher banal und nur für kämpferisch gesinnte Antiidealisten aufregend. Die wirklich interessanten Fragen beginnen erst jenseits solcher Grundannahmen. Für unseren Zusammenhang ist es die Frage, wie Informationen zwischen kognitiven

4 Latour geht dabei so weit, zu behaupten, dass wir eigentlich nicht denken, sondern metaphorisch gesprochen nur »schreiben«, also »an action which involves working with inscriptions«. Und weiter: »an action that is practiced through talking to other people who likewise write, inscribe, talk« → »an action that convinces or fails to convince with inscriptions which are made to speak, to write, and to be read« (LATOUR 1988: 218). Ob wir wirklich nicht denken oder Ideen haben, ist im theoretischen, auf Abstraktion beruhenden Sinn zu bezweifeln (ist eine Frage der Modellierung, bei der auch neuronale Inskriptionen anzusetzen wären).

Systemen wandern können (eine Frage, die auch den radikalen Konstruktivismus umgetrieben hat), welche Rolle dabei intermediäre Instrumente spielen und wie die damit zusammenhängenden Vorgänge funktionieren. Notiertheit ist die Elementarbedingung der Existenz von Texten im kommunikativen Interaktionsraum. De Beaugrande und Dressler nennen es »kommunikative Okkurrenz«. D. h. Texten eignet das Merkmal des »Vorkommens« in der materiellen Welt. Texte sind okkurent (DE BEAUGRANDE/DRESSLER 1981: 3; siehe auch Problem 21: Transnotation).

Den Notationszwang hat man schon länger gesehen, wenn auch nicht mit der nötigen begrifflichen Klarheit. So schreibt etwa Helmut Willke 2004: »Daten müssen in irgendeiner Form codiert sein, damit sie existent werden und ›gelesen‹ werden können. Bemerkenswerterweise sind die für Menschen intelligiblen Formen der Codierung extrem beschränkt, nämlich auf genau drei Möglichkeiten: Zahlen, Sprache oder Texte und Bilder. Was nicht in diese Codierungsformen gepresst werden kann, ist als Datum nicht existent« (WILLKE 2004: 29).

In unserem Zusammenhang muss dies korrigiert und präzisiert werden. Wir sagen: Mitteilungen müssen in irgendeiner Form encodiert sein, damit sie in die Interaktion sinnvoll eingespeist werden können (Adpragmatisierung). Menschen bedürfen dazu bestimmter InformationsCodes (künstliche oder kulturell verankerte Zeichenkonventionen) als textproduktionsregulierender Systeme. Diese sind Voraussetzung jeglicher Textproduktion und Kommunikation. Für die körperexterne Notation von Texten (Zeichenkomplexen) benötigen die Menschen spezielle NotationsCodes, die an jene verschiedenen medialen Kanäle angepasst sind, die die menschlichen Sinne erreichen (akustisch, visuell, haptisch, olfaktorisch, gustatorisch). Am weitesten verbreitet sind die alphanumerischen Notationscodes (etwa die mediterranen phonetischen Schriften), aber auch die asiatischen Ganzwort-Schriften sowie der BildCode und andere kulturelle SymbolCodes (SCHNOTZ 1990: 177–180). Mit ihrer Hilfe werden Texte physisch präsent und kommunikabel. Man kann aber auch Einzelsymbole aus NotationsCodes in die Kommunikation einspeisen. Was nicht körperextern annotiert ist, kann nicht als Mitteilung kommuniziert werden.

In den Künsten wird der Begriff der »Notation« gewöhnlich für Transcriptverfahren benutzt, bei denen es um die partiturartige Aufzeichnung »notierter und wiederholbarer Vorgehensweisen« geht (APPELT 2008: 16). Das Modell liefert die Musiknotation, auch wenn es inzwischen unter Musiktheoretikern die Auffassung gibt, dass auch bei der Musik »die aktuelle Notation ein Text und nicht eine bloße Anweisung sei« (LAMMERT

2008: 39). Das traditionelle Musiknotationssystem stand in einem gewissen Sinn auch Pate für Nelson Goodmans Notationskonzept mit seinen streng algorithmisch gedachten Regeln, das auf die in unserem Zusammenhang wichtige Primärnotation (wenn auch anders gedacht) ausgreift (GOODMAN 1997 [1968]; KRÄMER 2003: 163; vgl. BÖHNKE 2004: 175; zum Begriff des Algorithmus in Notationssystemen vgl. WEIBEL 2008: 34f.).

Im Unterschied zu Goodmans sehr spezieller Notationsvorstellung ist im Rahmen der Bildtheorie allerdings unter *Notation* ganz allgemein jeder Vorgang der Inskription als Schaffen und Erkennbarmachen eines außerhalb des Körpers manifestierten Zeichenkomplexes, also eines Textes, zu verstehen (KNAPE 2010: 83).<sup>5</sup> Kurz: Der Text muss für Menschen wahrnehmbar medialisiert sein. Informationalität des Bildes, zu der wir hier ›Bedeutung‹ des Bildes sagen, kann nur im Zusammenspiel und Reaktionsverbund von kognitiven Systemen (Menschen) und Angeboten körperexternalisierter semiotischer Manifestationen (z. B. Bildern) enkodiert und dekodiert werden. Informationen werden im Allgemeinen materiell mittels Zeichen zugänglich gemacht, die von den menschlichen fünf Sinnen wahrnehmbar sind. In der Forschung ist inzwischen die Frage »Wie kann mit welchen Medien notiert werden?« durchaus auch mit Ausblick auf die hier von uns herausgestellte Primärnotation diskussionswürdig geworden (LAMMERT 2008: 39). Produktionstheoretisch gesehen sprechen wir beim Textnotieren ersten Grades vom Aufzeichnen, Aufschreiben, Klangformulieren, Signalisieren, Sprechen, Malen, Fotografieren usw. »Eine Schrift ist somit ein notationales Medium«, sagt die Medienphilosophin Sybille Krämer (KRÄMER 2003: 163; vgl. BÖHNKE 2004: 175). Der Begriff der Notation oder auch einfach des Notierens ist also nicht etwa nur auf schriftliche Medialisierung von lautsprachlichem Text bezogen, sondern intersemiotisch auf Texte eines jeden beliebigen semiotischen Systems (KNAPE 2008: 896f.). Auch Apparate können notieren.

Schon am Beginn der abendländischen Philosophie ist die Notiertheit bzw. der damit verbundene Text-Medium-Zusammenhang ein erstrangiges Thema. In Platons spätem Dialog *Phaidros* (274b–277a) aus dem 4. Jh. v.u.Z. findet sich die berühmte platonische Schriftkritik, die die neue Qualität

5 Man kann sagen, dass der Gebrauch des Begriffs ›Notation‹ in der heutigen Theoriebildung »von einer Spannung der Gegensätze gekennzeichnet« ist. »Einerseits gilt die enge Auffassung der musikalischen Notation als Partitur. Andererseits gilt der umfangreichere Begriff der Notation als Aufzeichnung« (WEIBEL 2008: 32).

der seit dem 8. Jh. bekannten Schriftnotation von lautsprachlichen Texten im Mittelmeerraum problematisiert und kritisiert (zusammenfassend KNAPE 2015: 120–122).

Für unseren Zusammenhang ist aber die Frage der Beschaffenheit der NotationsCodes oder NotationsArsenale, die für die Körperexteriorisierung von Bildern herangezogen werden, von erstrangiger Bedeutung. Zunächst einmal muss zwischen InformationsCodes und NotationsCodes oder den Notationsform-Arsenalen begrifflich unterschieden werden (KNAPE 2024: 86). Die Informationscodes sind die Sprache oder die von uns auch noch systemisch gelernten Codes (z. B. Bildwissen), die für uns als Zeichenmaterialbasis und Hintergrundwissen für die Textproduktion fungieren. Sie sind letztlich abstrakte, kognitiv verankerte Systeme. Das Paradebeispiel geben die verschiedenen menschlichen Lautsprachen ab. Wenn ihr Systemwissen aktualisiert wird, also in Form von Texten hervortritt, dann brauchen wir für die Zeichen NotationsCodes. Beim Bild sind das die kulturell verankerten, intersubjektiv verständlichen Bildzeichen. Bei der Lautsprache aber haben wir, da sie digital funktioniert, mehrere Notationsmöglichkeiten, entweder über den akustischen Kanal (Sprechen) oder über den optischen (Schreiben, Gebärden usw.; KÜMMEL 1969). Das ist eine Besonderheit bei der Umsetzung des Informationscode Lautsprache in seine Notationsformen bei der Textproduktion: Es gibt einen InformationsCode = Lautsprache, aber mehrere NotationsCodes. Prominent ist hier natürlich die Optimierung lautsprachlicher Texte mittels schriftlicher Notationscodes oder mittels Gebärden. Beim Bild als visuellem Textformat im strengen Sinn gibt es auch einen InformationsCode = BildCode, aber nur einen einzigen NotationsCoder der Bildnotationszeichen. Das Besondere: Notationszeichen fallen hier mit den Informationszeichen ausdrucksseitig zusammen.

Wenn also die Philosophin Sibylle Krämer behauptet, auch Schrift könne Bild sein, dann trägt das auf diese Weise formuliert nur zur weiteren Verwirrung bei. »Schriftbildlichkeit« (KRÄMER 2003), wie sie es nennt, ist eine *contradictio in adiecto*. Auch andere behaupten Bild, Schrift und Zahl seien nicht »strikt voneinander zu trennen« (HESSLER/MERSCH 2009: 10). Demgegenüber ist geltend zu machen, dass Schriftnotate und Bildnotate zwei differente Visualisierungsformate ergeben und sich als solche ganz klar diskret verhalten (KNAPE 2023b: 20f.). Krämer spricht eben nicht – ohne uns das zu erklären – von den Verhältnissen auf der semiotischen Bildtextebene, sondern (mal wieder) von kognitiven Verarbeitungszusammenhängen im Menschen, die auf einer ganz anderen Ebene liegen. Und

da kann es bei der kognitiven Verarbeitung von Schriftwahrnehmungen in der Tat zu bestimmten Imaginationüberschneidungen kommen, die beim Anschauen von Schriftseiten als analoge Imaginationen aufgenommen werden. Das ist aber was für Neurologen.

#### 2.14 *Vierzehntes Problem: Pragmatik*

Jedes Bild unterliegt einer pragmatischen Grundorientierung der Kommunikation. Bilder existieren aus pragmatischen Gründen (BELTING 1998). Die Bildpragmatik befasst sich mit den interaktionalen Einbettungsbedingungen des Bildes, so etwa auch mit dem kommunikativen Status, dem Verstehens-Frame und den Gattungszuordnungen. Bilder sind semantische Angebote an Bildverwender, die zugleich Settingswissens- und Bildwissensinhaber sein müssen. Wenn Bildbetrachter dieses Wissen nicht einsetzen können, bleiben Bilder in vielerlei Hinsicht bedeutungslos.

Produktion und Rezeption des Bildes unterliegen in der Normal- oder Standardkommunikation den Bedingungen der Standard-Frame-Prämisse. Was heißt das? Nach der Prämisse akzeptieren die Bildermacher und die Bilderdeuter die Alltagswahrnehmungen der Menschen als Norm und leiten daraus ein Realismuspostulat für Enkodierung und Dekodierung von Informationen auf der Textebene, sprich: Bildzeichenebene, ab (WEDEWER 1985: 161–170). Damit regulieren Sehgewohnheiten die Bildkonstruktion und Bilddeutung (KNAPE 2013a: 430; zum »naturalistischen Identifikationsmodus« siehe DESCOLA 2023: 621–632; vgl. WEDEWER 1985: 191). Die Standard-Frame-Prämisse ist eine semantische Prämisse, die davon ausgeht, dass im Bild zwar »virtuelle« (KNAPE 1997, 2019: 43–45), aber informationell an der *actual world* orientierte und darum verstehbare *possible worlds* aus Bildzeichen (und eventuell weiteren kulturellen Symbolen) dargestellt werden. Das ist die Standard-Realismus-Prämisse der Bildtheorie. Damit ist nicht gesagt, dass es nur einen kommunikativen Funktionstyp des Bildes gibt (siehe Problem 6: Funktionstypen, Format- und Gattungsfragen).

Produktion und Rezeption des Bildes werden gemäß Standard-Frame-Prämisse immer durch die *Kontextbedingung* reguliert, die sich als Implikation in den semantischen Angeboten des Bildes niederschlägt und seine Verwendung in bestimmten kommunikativen Settings als Bedeutungsträger funktional macht (siehe Problem 9: Kommunikabilität). Bilder entstehen nur (wenn es kein zufälliges Gekritzeln sein soll) in Hinblick auf kommunikative

*Einbettungen* in bestimmten Settings. Und nur in Hinblick darauf können sie auch in ihrer Mitteilungsfülle interpretiert werden. Der Bildtext ist also immer adpragmatisiert zu sehen, d. h. verbunden mit kommunikativen Interaktionskontexten, die Mitteilungsziele definieren und hermeneutische Schlüssel bereithalten, sowie verbunden mit kognitiven Schnittstellen der Bildwissensinhaber. Bildkommunikation kann misslingen, wenn das nicht der Fall ist. Zu den wesentlichen Einbettungen von Bildern gehören rhetorische Kommunikationsprozesse, von denen später noch gesondert zu reden ist.

Die Bildsemantik kann aber auch durch andere rahmende Prämissen reguliert sein. Bildverstehen ist nicht auf die Realismusprämisse festgelegt. Den hermeneutischen Schlüssel liefern oft die Frames der Kunst, Ästhetik, Philosophie, Religion, bestimmter Rituale oder sonstiger kultureller Praktiken.

### 2.15 Fünfzehntes Problem: Semiotizität

Ein Bild ist ein geordneter und begrenzter Bildzeichenkomplex in kommunikativer Absicht, der zweidimensional medialisiert ist. Folglich ist nicht jedes optische Artefakt ein Bild (KNAPE 2023b). So gibt es etwa viele Gemälde, die keine Bilder sind (z. B. abstrakte Gemälde oder reine Farb- oder Ornamentkompositionen), weil sie nicht in ihrer Grundstruktur auf Basis des BildCodes komponiert sind. Da Bildzeichen aber semantisch unterdeterminiert sind, (SUSANKA 2015: 154–158) entstehen regelmäßig Hybridvisualisierungen mit Codemix unter Hinzunahmen von diversen weiteren visuellen Elementen aus diversen kulturellen SymbolCodes (vor allem die lautsprachebezogene Schrift).

Als Zeichengestell muss das Bild gemäß Semiotizitäts-Axiom auf Zeichenebene theoretisiert werden (Überblick bei NÖTH 2000: 471–480). Sehr oft wird jedoch die körperextern medialisierte Zeichenebene mit der theoretischen Ebene kognitiver Verarbeitung der Zeichen im Menschen vermischt oder verwechselt. Die neuronalen, kognitiven und psychologischen Verarbeitungsprozesse als solche fallen jedoch nicht ins Gebiet der Bildtheorie (siehe Problem 4: Bindestrich-Theorien). Bemerkenswert ist allerdings die Schnittstelle (BELTING 1998). Was sich im Menschen bei der Rezeption und mentalen Verarbeitung abspielt sollte man nicht als Bildgebung bezeichnen, sondern unterscheidend als körperinterne reaktive *Imaginationen* (die in unserem Denken neuronal vielleicht gar keine wirklichen ›Bilder‹ sind;

niemand kann es wissen, weil es niemand sehen kann). Ich spreche ausdrücklich von kognitiven Imaginationen, um mit diesem Differenzbegriff auf den Unterschied von medialisiertem Bildtext und Kognitionen als Informationsverarbeitungsprozessen hinzuweisen. Notabene: Bei der Bildzeichenverarbeitung führt die spontane Verwechslung von Bildwahrnehmung mit der natürlichen Sinneswahrnehmung äußerer Gestalten oft zu gewissen Stimulus-Response-Situationen: phobische Ablehnung, Schrecken, Rührung, Erregungen aller Art usw. (TVERSKY 1977; MARKMAN/GENTNER 1993). Insbesondere die Kritiker einer modernen semiotischen Bildtheorie vermischen gern die beiden theoretischen Perspektiven (Bild versus Bildverarbeitung; so etwa BACHMANN-MEDICK 2007: 343f.). Wissenschaftlich gesehen kann jedoch nur die Bildsemiotik produktions-theoretisch die Enkodierung, also den Aufbau von Bildsemantiken als Vorgang, sowie rezeptions-theoretisch die Dekodierungsmöglichkeit von Bildern plausibel erklären.

Gemäß semiotischer Zwei-Sphären-Theorie (Lebenswelt nach Husserl versus Zeichenwelt) bildet die Ebene der Zeichen ein eigenes, theoretisch abstrahiertes Universum. Ihre Manifestationen (Symbole und Texte) sind auf einer eigenen analytischen Ebene zu verhandeln. Bildzeichen sind eigenständige, von ihren gestaltgebenden Realobjekten erlöste Symbole, die einem autonomen Code zugehören, ähnlich wie die Wörter des VerbalCodes. Wie jedes Zeichen, steht auch das Bildzeichen als solches nicht referenziell für ein konkretes, sichtbares Objekt der physikalischen Welt, sondern als abstrahierte Größe für eine ganze Gegenstandsklasse im Bewusstsein. Die Bildzeichen 🌳 und 🐎 wurden an prominenter Stelle vom großen ersten Beweger moderner Zeichentheorie, Ferdinand de Saussure, als Paradigmen für Bildzeichen eingeführt. Sie stehen als *types* oder *Bildzeichen* in der Bildkommunikation für alle denkbaren Bäume und Pferde (LANGER 1984 [1942]: 76–81). Ob sie im konkreten Anwendungsfall auch als »Abbilder« auf ein jeweils ganz bestimmtes Außenwelt-Objekt referenzialisieren, legen erst der Kontext sowie gewisse *token*-Ausprägungen der *types* im Zeichengestell fest.

Man kann sich mit Bildern per »Deixis am Phantasma« (BÜHLER 1934: 123f.) verständigen, also nur über das durch Bildtext herbeigeführte Auslösen gedanklicher Vorstellungen (Imaginationen) innerhalb des kotextlichen, bühlerschen »Zeigfelds«, die im Moment der Bildbetrachtung evoziert werden. Das gilt für Einbettungen des Bildgebrauchs in nahezu alle Kommunikationszusammenhänge, freilich nur im Modus des reinen Vor-Augen-Stellens in räumlich geordneten Zeichenkonstellationen, weil der BildCode auch keine »Zeigwörter« kennt, wie sie Bühler nennt.

## 2.16 Sechzehntes Problem: Sprachanalogie

Eine Sprache (*langue*) ist ein System von Zeichen mit einem bestimmten Zeichenvorrat (Lexikon) mit entsprechendem Ausdruckssystem (etwa einem akustischen Phonemsystem, optischen Graphem- oder Gestaltformensystem) sowie einem Regelwerk für die Strukturvarianten der Zeichen (Morphologie der Einzelzeichen) und ihrer syntagmatischen Kombinationsmöglichkeiten. Die Sprache liefert nur das Baumaterial für die Konstruktion von Text->Gebäuden<, die Semantik konstituieren.

Das bildtheoretische Konzept der ›Sprachanalogie‹ sieht einerseits vor, dass Bilder gewisse kommunikative Leistungen vollbringen können, wie sie auch Texte mit Hilfe von Lautsprachen (manifestiert in Gestalt von Verbaltexten) vollziehen können (DESCOLA 2023: 670–680, der Sprache und Text meist leider nicht als Kategorien klar trennt). Andererseits sieht das Konzept vor, dass es einen Bildzeichenkode gibt, der auf abstrakter Ebene einige strukturelle Gemeinsamkeiten mit anderen Kodes in der Welt semiotischer Systeme aufweist, z. B. Konventionalität des Zeichenfundus unter den CodeBenutzern (NÖTH 2000: 341). Es sieht nicht vor, dass diese Systeme zwingend ähnliche Grammatiken (Richtigkeits-Regelwerke) haben müssen.

Eine »Sprache ist ein System von Zeichen, die Ideen ausdrücken«, heißt es 1915 beim wichtigsten Zeichen- und Sprachtheoretiker des 20. Jahrhunderts, Ferdinand de Saussure, wenn eine Sprache empirisch zur Konstruktion von Texten verwendet wird (DE SAUSSURE 1967 [1916]: 19). Damit hebt De Saussure die Bedeutungsdimension alles Semiotischen hervor. Ob solch ein Zeichensystem ›Sprache‹ mit einem ›Code‹ im weitgefassten Verständnis begrifflich gleichgesetzt werden kann oder ob beide Begriffe differente Perspektiven aufmachen, muss diskutiert werden. Ich nehme den Terminus ›Code‹ hier als Oberbegriff für Kodierungssysteme, unter den auch die ›Sprachen‹ fallen.

Was all das aber im Detail heißt, wird in der Forschung vielfältig und widersprüchlich erörtert (NÖTH 2000: 478f.). Manche Bildtheoretiker verstehen gar nicht, was Sprache als systemische Größe überhaupt ist, etwa im Unterschied zur Textualität (BOEHM 1978: 452–455), so dass Wollheim 1996 den dubiosen Satz formulieren kann: »painting can function sometimes like language, sometimes not like language« (WOLLHEIM 1996: 37). Hier werden die Ebenen von Sprache (*langue*) und Text (Ebene der *parole*) verwechselt bzw. gar nicht gesehen. Manche lehnen das Lautsprachen-Analogietheorem rundweg ab (BOEHM 1978: 447f., 458, 463; vgl. WEDEWER 1985: 31,

39). Winfried Nöth will sich im Jahr 2000 nicht festlegen und spricht nur vorsichtig über »die Metapher von der ›Sprache des Bildes««. Wenn man die gerade zitierte Definition von De Saussure als Ausgangspunkt nimmt und zugesteht, dass Lautsprache und Bildsprache als Systeme zwar gemeinsame Merkmalschnittmengen haben, aber nicht identisch sind und die Lautsprachenanalogie ihre Grenzen hat, dann ließe sich der Begriff ›Bildsprache‹ vielleicht auf einer bestimmten Abstraktionshöhe rechtfertigen.

Im Unterschied zum ›Text‹ ist die ›Sprache‹ als System kommunikations-situationserlöst (WEDEWER 1985: 121) und lediglich der Baustein- und Korrektheitsregel-Lieferant für die Konstruktion von Text. Gemäß Definition ist ein Bild ein Text und damit ein begrenzter, geordneter und stillstehender Bildzeichenkomplex in kommunikativer Absicht. Für seine Konstruktion braucht der Bildermacher Bausteine. Sie liefert das Bildwissen, in dem das systemische Wissen des BildCodes eines Bildverwendungskreises verankert ist. Man kann daher den BildCode mit einer gewissen Berechtigung auch ›Bildsprache‹ nennen.

Die Modellierung könnte dann in Analogie zu den für die Lautsprache gefundenen Systemmerkmalen erfolgen. Man würde dabei von den folgenden sechs systemisch abstrahierten Merkmalen ausgehen (vgl. zum Folgenden: WUNDERLI 1981: 33; NÖTH 2000: 73): 1. *Arbitrarität* der Zeichen (siehe dazu unten unter Problem 23: Zeichen); 2. BildCode als *Institution* der BildCode-Verwendungsgemeinschaft, die damit über ein regional oder epochal begrenztes Zeichenrepertoire verfügt; 3. *Immutabilität*, d. h. Unveränderlichkeit der Bildzeichen für das Individuum (wenn es den Code benutzt); 4. *Fehlen jeglicher im Voraus festgelegter Einheiten*, d. h. die Bildzeichen sind in ihrer Ausdrucks- und Inhaltsform durch nichts als durch das Bildwissenssystem bestimmt; 5. *Produktivität*, d. h. es gibt unbegrenzte Möglichkeiten der Produktion von Texten mit dem regional begrenzten Zeicheninventar;<sup>6</sup> 6. *mediale Manifestation* der Zeichen (Lautsprache akustisch,<sup>7</sup> Bildsprache optisch).

6 Begrenztheit des Bildzeicheninventars heißt, dass nicht alle Bildzeichen in allen Weltgegenden verstanden werden können (vgl. das Bildzeichen  $\alpha$  bei KNAPE 2016: 40–43).

7 Der NotationsCode ›Schrift‹ wird dabei als Sekundärphänomen der Lautsprachlichkeit (!) gesehen.

2.17 *Siebzehntes Problem: Still-Theorem*

Ein Visualisierungsformat ist nur dann ein ›Bild‹ im bildtheoretischen Sinn, wenn das Stillgestelltheits-Axiom gilt: Ein Bild ist ein *still*, kein ›Fließ-‹ oder ›Bewegtbild‹ (*movie*, *motion picture* oder *film*, die Bewegungswahrnehmungen ermöglichen). Die Bildzeichen sind also im visuellen Raum der Bildtextur stillgestellt. Insofern fehlt dem Bild auf Textebene und performativ auf Medialisierungsebene im Unterschied zum Film eine lineare Zeitachse. Um Missverständnissen zu begegnen, sei hier auch noch etwas eigentlich Selbstverständliches hinzugefügt: ›Bilder‹ sind keine Bilderfolgen, Bilderzyklen oder Bilderreihen. Semiotisch gesehen bedeutet all dies, dass ein Bild keine Zeitrelationen ausdrücken kann, sondern nur die räumliche Ausdehnung von Gestellkonstellationen der Bildzeichen im Raum. Was daraus bei der Bildwahrnehmung und mentalen Bildverarbeitung im Menschen wird, steht auf einem anderen Blatt.

2.18 *Achtzehntes Problem: Syntagma*

Gleich zu Beginn sei hervorgehoben, dass Bilder keine ›Syntax‹ haben, weil sie als nicht-lineare Zeichengestelle keine Sätze im linguistischen Sinn enthalten. Jedes Bild ist freilich eine Textkomposition.<sup>8</sup> Ein Bildtext ist gekennzeichnet von Kompositionalität (siehe Problem 20: Textualität). Erst der Verbund mehrerer Bildzeichen begründet den Tatbestand der Textualität. Ein alleinstehendes Bildzeichen oder Symbol ist kein Bild, sondern das was es ist. Ein Bild ist mithin ein eigengesetzliches »Symbolfeld« (BÜHLER 1934: 372), das in Form von Gestellen von Syntagmen organisiert ist, die in einer räumlichen Zusammenstellung verbundene Gestaltformationen darstellen (zum Bildraum als Ordnungsgefüge und entsprechende »Füge-techniken« im räumlichen Syntagma siehe WEDEWER 1985: 143–154; vgl. auch SCHAPIRO 1995). Ein Text funktioniert in seiner Organisation ganz individuell. Das gilt nicht in gleicher Weise für seine Bausteine, Einzelkomponenten oder die syntagmatischen Muster.

8 Grundsätzliches Unverständnis (insbesondere sprachlich-linguistischer Theorien und entsprechender Analogieverhältnisse) bei БОЕИМ 1978: 450–453.

Es gibt keine Bildsyntax, aber Bildsyntagmen, die Bedeutung organisieren. Da Bilder, wie jeder Text, in erster Linie Bedeutungsangebote liefern, achten Bildproduzent und Bildinterpret nicht nur auf die Auswahl der Einzelzeichen, sondern auch auf die semantischen Solidaritäten in den »syntagmatischen Beziehungen« (GADLER 1992: 152) der Bildelemente im Bedeutungsnetz des Textes: Welches Bildzeichen ist neben welchem Bildzeichen gestellt und welche Bedeutung ergibt sich sinnvollerweise aus genau diesem Zeichengestell? Die beim Bildinterpretieren im Moment der Wahrnehmung auftretende semantische »Erwartung« von bestimmten Bildzeichen nach bildwissensmäßig gespeicherten Mustern werden nach linguistischem Vorbild »Kollokationen und die Beziehung als *Kollokabilität* bezeichnet« (GADLER 1992: 152f.). Treten Erwartungsbrüche in den üblichen, semantisch bedingten Verbindungen als nicht mehr ohne weitere verstehbare Zeichenverbindungen auf, stellt sich dem Betrachter die Frage nach dem geltenden ZeichenCode (normaler BildCode oder andere SymbolCodes?) bzw. der im semantischen Bruch zum Ausdruck kommenden Intention.

Unter den Bedingungen der Standard-Frame-Prämisse (siehe Problem 14: Pragmatik) werden die Seherfahrungen, Sehgewohnheiten und kulturspezifischen Wahrnehmungsmodelle zu Deutungsregulativen, die sich in den Codes und damit im Bildwissen und Weltwissen niederschlagen (BERGER 1974; SACHS-HOMBACH 2007: 54–57; vgl. zum Problem von Wahrnehmung und »Sichtweise« im Bild WIESING 2007). Sie betreffen die Bildelemente (motivierte Zeichen) und die Zeichengestelle im Text (Syntagmen). Gemäß dieser Realismusprämisse sind vor allem Perspektive, Ausrichtung der Objekte (Vektorialität nach oben, unten usw.) sowie alle sonstigen Positionen im Raum nach den Modellen der alltäglichen Seherfahrungen des Menschen reguliert (TUCHOLSKI 1971). Sowie aber die Realismusprämisse nicht mehr gilt, kommen weitere Symbol- und Darstellungsregulative ins Spiel (ästhetisch, religiös, ideologisch etc. motivierte; KNAPE 2010: 85).

Weil es im Bild keine »Sätze« gibt, wie in den linear verlaufenden lautsprachlichen Texten, kann es keine Syntax im Rahmen des BildCodes geben. Für den Aufbau von Bildern müssen sich Bildermacher also nicht von einer systemisch geforderten »formalen Syntax« nach linguistischem Vorbild abhängig machen, weil im Bildtext nur semantisch organisierte Bildzeichen->Cluster« auftreten, deren Konstruktionslogik nicht mit der in den »Sätzen« der Lautsprachentexte vergleichbar ist (vgl. NÖTH 2000: 479f.). Und formale Äquivalente zu den linear und logisch beigeordneten »Prädikatsphrasen« des lautsprachlichen Textes gibt es auch nicht. Im

Bild sind ›Prädikate‹ Gestaltelemente als Bestandteil komplexer, in sich untergliederter Bildzeichen (Gestaltentitäten im gestalttheoretischen Sinn; näheres WEDEWER 1985: 135, 142f.; SACHS-HOMBACH 2007: 57–59).

Konventionelle Bildordnungsmodelle sind vielfältig, je nach Bildsprachenverwendergruppe. Das basale, semantisch motivierte Bildordnungsmodell Europas ist gemäß Standard-Frame-Prämisse naturalistisch (*ordo naturalis*) und gründet sich auf die in der Alltagswahrnehmung gelernten Ordnungsstrukturen der physikalisch sichtbaren Dinge der Welt bzw. der sichtbaren Wirklichkeit außerhalb des Bildes, die naturalistische Ordnungsanalogien im Bild nach sich ziehen (›Logik des Sichtbaren« [Konrad Fiedler]; BOEHM 1978: 446). Die Bildzeichen können aber auch nach einem künstlichen Ordnungsmodell (*ordo artificialis*) im Raum angeordnet werden. In diesem Fall regieren beim Bildermachen die schon genannten religiös, mythisch, traditionalistisch oder ästhetisch motivierten Ordnungsregulative (als Konventionen) das Syntagma und determinieren die Bildtextordnung. Es versteht sich, dass solche Ordnungsmodelle (1) in der Standardkommunikation bei den Rezipienten nach der Common-Ground-Bedingung bekannt oder als Abweichung erkannt sein müssen und dass demgegenüber (2) in der Lizenzkommunikation (etwa im Kunstfall) auch beim Syntagma jede Form von kreativer Devianz, Übertretung und Abweichungen vom Analogieansatz möglich ist.

Bei der Konstruktion von Bildtexten ziehen die Bildermacher solche Regulative heran, die ihren Mitteilungsententionen gerecht werden und zugleich die Verständigung mit den Bildbetrachtern sichern. Kurz: Produktionstheoretisch gesehen ordnet der Bildermacher die Bildzeichen nach semantischen Kriterien (Aufbau einer sinnvollen Struktur) gemäß seiner Kommunikationsabsichten als Gestell im Bildraum. Um die Verstehbarkeit zu sichern, muss er dabei auf die Frames eines gemeinsamen Bildsyntagmawissens rekurrieren. Die Geschichte der Geltung der Perspektive im Bild kann das verdeutlichen (DESCOLA 2023: 67ff. und 623ff.). Die radikale Perspektivenbedingung (nach dem Vorbild der physikalischen Optik) setzten erst die Bildermacher im Europa der Renaissance nach und nach durch. Unter den erwähnten Bedingungen der Standard-Frame-Prämisse gibt es insgesamt verschiedene standardisierte Bildgattungsrepertoires und -ordnungen, die aber nie zwingend sind. Kommen jedoch im Fall des lizenzierten Kommunikationsstatus weitere Symbol- und Darstellungsregulative ins Spiel (ästhetisch, religiös, ideologisch etc. motivierte), ändern sich die Regeln. Da kann dann auch etwas möglich werden, das nach den alltäglichen Sehordnungen und Regeln der Physik unmöglich, in der *pos-*

*sible world* eines anderen semiotischen Ordnungssystems aber erlaubt ist (KNAPE 2024: 108–110). Maler wie Salvador Dali oder René Magritte – um nur diese zu nennen – haben von diesen Möglichkeiten ausführlich Gebrauch gemacht (Beispiele bei KNAPE 2013a: 428f.).

Wie gesagt, bei Bildtexten fehlt zwar die Strukturebene des ›Satzes‹, mit der Folge, dass es auch keine Syntax gibt, doch es gibt in jedem Bild Syntagmen (bedeutungsrelevante, sinnvolle Verbindungen bei den Gestellzusammenstellungen). Es gibt auch eine semantische Segmentierungshierarchie. Deren Einheiten sind (1) Bildzeichen (siehe Problem 23: Zeichen), (2) semantisch organisierte ›Cluster‹ im Bild (BALDRY/THIBAUT 2006: 31; KNAPE 2024: 92f.) sowie (3) das Bild in toto mit seiner übersummativen Semantik. Diese Textbedingungen sind abgestimmt auf die bildverarbeitenden Kognitionen im Menschen, insbesondere mit seinen Segmentierungsleistungen bei der Bildinformationsverarbeitung. Menschen können gemäß ihrer Lerngeschichten im Bild, wie in allen Texten, semantische Einheiten als vom Umfeld abgehoben und im Verbund zugleich als höher organisierte Bedeutungseinheiten erkennen.

Große theoretische Verwirrung hat in diesem Zusammenhang die von Nelson Goodman aufgebrachte These der angeblichen Existenz von »Dichte« (GOODMAN 1997 [1968]: 133f., 154–156), einer »unstillbaren Übergänglichkeit« (BOEHM 1978: 462f.) im Bild gestiftet. Dichte soll die Unmöglichkeit bezeichnen, eine Bildtextur auf diskrete Einheiten hin zu untersuchen. Bei Bildern sei es so, dass man zwischen zwei als signifikant angenommenen Punkten noch weitere dazwischen liegende annehmen müsse (vgl. BÖHNKE 2004, 173). Diese haltlose These ist ohne Einbezug moderner linguistischer Erkenntnisse aufgestellt und in ihrer Modellierung undurchdacht. Dabei handelt es sich in Wirklichkeit nur um ein Problem des Zusammenspiels von Wahrnehmung und Bildwissen.

Produktionstheoretisch gesehen konstruieren Bildermacher Bilder mit Hilfe semantisch segmentierter Einheiten: Der Bildermacher verwendet das als *type* definierte und segmentierte Zeichen 🐾, wenn er die Information ›Pferd‹ geben will. Rezeptionstheoretisch wird das später nur dann zum Problem, wenn Perzipienten in bestimmten Gegenden der Welt, wo man keine Pferde kennt, auch das entsprechende Zeichen nicht gelernt haben. Die sogenannte ›dichte‹, Übergangslos-fließende Wahrnehmung von Inskriptionen (wie wir sie auch beim Lautfluss akustischer sprachlicher Wahrnehmungen haben!) ist also lediglich ein Problem der Codekenntnis (siehe Problem 16: Sprachanalogie). Auch hier muss wieder deutlich die

semiotische Ebene mit ihren Fragen der Zeichendiskretion von der Ebene der Verarbeitungskompetenz der Menschen unterschieden werden. Auf Ebene des Bildes als Konstruktion gibt es keine syntagmatische (und schon gar nicht eine ›syntaktische‹, wie gesagt wird) oder semantische Dichte. Diese könnte man höchstens in Visualisierungsformaten annehmen, die keine ›Bilder‹ sind, etwa bei abstrakten Gemälden oder reinen Farbkompositionen, nicht jedoch beim ›Bild‹.

Wieso ist die semantische Struktur der Bildkomposition nicht mehr und nicht weniger von ›Dichte‹ gekennzeichnet als die der linearen lautsprachlichen Texte? Diese Frage beantworten die Dichte-Postulierer nicht. Wer den phonetischen Lautfluss einer Äußerung in einer fremden Sprache hört, die er nicht kennt, kann nicht segmentieren, erlebt nur Geräuschabläufe und einen für ihn amorphen Lautstrom. Erst wenn ein Hörer die linguistischen Einheiten einer Fremdsprache gelernt hat, kann er plötzlich beim Hörerlebnis bedeutungstragende Segmente aus dem unsortiert strömenden Geräusch- und Lautfrequenzüberfluss herausfiltern. Gleiches gilt für die optische Wahrnehmung von semantik-tragenden Segmenten im Bild, wo bedeutungstragende Einheiten ebenfalls mithilfe des lebenslang erlernten Bildwissens herausgefiltert werden müssen, um irgendetwas von der Bildinformation zu verstehen. Wenn Menschen nicht lernen könnten, was semantiktragende Gestalt-Einheiten im Bild sind, könnten sie keine Bilder deuten. Sie könnten auch keine Verbindungen zwischen den semantischen Segmenten herstellen und semantische Zusammenhänge im Bildtext erkennen. Das Lernen entsprechender Zeichen basiert – wie gesagt – auf Analogieschlüssen aus den natürlichen Gestaltwahrnehmungen der physikalischen Außenwelt, die nach und nach bei den Menschen zum mentalen Bildlexikon als Teil des Bildwissens werden.

Eine akustische Laut-›Dichte‹ gibt es auch beim Hören menschlicher Verbaltexte. ›Erst die Differenzierung von Getrennt- und Zusammenschreibung erzeugt das ›Wort‹ als grammatische Kategorie, und mit dem ›Satz‹ verhält es sich entsprechend« (STETTER 1997: 10). Segmentierung ist die entscheidende kognitive Leistung beim Identifizieren bedeutungstragender Gestalteinheiten im undifferenziert ankommenden Wahrnehmungsfluss. Wie komplexe Bedeutungszusammenhänge in den Verbaltexten ermittelt werden, erklärt der Sprachwissenschaftler Helmut Lüdtke. Seine Erkenntnisse lassen sich auch auf die Bildverarbeitung übertragen: Aufgrund des sprachökonomischen Optimierungsstrebens kommt es in Sprechergemeinschaften oder in Bildzeichenbenutzergemeinschaften dazu,

»dass ständig in der morphologischen Verkettung benachbarte Einheiten allmählich miteinander zu neuen Einheiten verschmelzen« (LÜDTKE 1980: 16f.). Somit »kann sich auch die durchschnittliche Größe der semantaktischen Einheiten, d. h. die ihnen zugeordnete artikulatorische *target*-Menge, nicht irreversibel ändern«. Dem Adressaten, also dem wahrnehmenden Textrezipienten (welcher Zeichensorte auch immer) kommt vor diesem Hintergrund eine wichtige Rolle zu. Er ist nämlich verantwortlich für den genannten »*Verschmelzungsprozess*, indem er bei der Verarbeitung des ankommenden Sprachschalles – oder beim Bild des ankommenden optischen Materials – dieses Material »*segmentiert*, d. h. analysierend in Einheiten zerlegt«. Diese Tätigkeit des Hörers von Lauttexten enthält »ein begrenztes Maß an Willkür«, insofern nämlich, als er etwa bei lautsprachlichen Texten »in zahllosen Grenzfällen (wie z. B. dt. *Rätsel*, fr. *aujord'hui*, e. *cupboard*) die fakultative Segmentierung vornehmen oder unterlassen kann«. Der »*homo loquens*« oder eben auch der *homo videns* »ist nicht nur sowohl Hörer als auch Sprecher«, sondern »außerdem Hör/Sprech-Koordinator« bzw. im Bildfall auch Seher und Bildermacher bzw. ein entsprechender Koordinator. Für sprachtextliches Erleben gilt (und Analoges können wir für bildtextliches Erleben postulieren): Indem der Zeichenbenutzer

»sein eigenes Sprechen kinästhetisch, als Bewegungsgefühl, wie auch auditiv, mit dem Ohr, wahrnimmt, ist er in der Lage, zwischen zeitlich abgrenzbaren Teilmengen von phonotorisch-artikulatorischen Bewegungs- und entsprechenden Teilmengen von Gehörseindrücken feste Korrelationen herzustellen und die korrelierten Teilmengen als abgegrenzte Einheiten sich bewusst zu machen. Die *Vorstellung* von den sprachlichen Einheiten beruht auf diesem – als Segmentierung bezeichneten – Koordinationsvorgang, der vom sprachlich reflektierenden Individuum ständig neu vollzogen wird und nach Belieben in sein Bewusstsein gehoben werden kann« (LÜDTKE 1980: 16f.).

In bildtheoretischer Analogie geht es allerdings nicht um »zeitlich abgrenzbare Teilmengen von phonotorisch-artikulatorischen« Gehörseindrücken, sondern um das Wahrnehmen optischer, räumlich benachbarter Teilmengen von pictonotorisch-visualisierten Scheindrücken.

## 2.19 Neunzehntes Problem: Terminologie

Es herrscht ein terminologischer Pauperismus in der Bildtheorie. Wenn die Bildtheorie als »wissenschaftliche« Theorie unter vergleichbaren The-

orien Bestand haben soll, dann muss sie als wesentliches Erfordernis der Wissenschaftlichkeit auch auf eine klare, diskrete und wohldefinierte Fachterminologie in ihrer Fachsprache achten. Ein solches Bemühen wird oft unter Jargon-Verdacht gestellt, dem dann ein naives Bestehen auf umgangssprachlicher, angeblich besser verständlicher Begrifflichkeit entgegengesetzt wird. In Wirklichkeit handelt es sich dabei meist um Theoriefeindlichkeit und inhaltliches Unvermögen, Klarheit zu gewinnen. Natürlich kann man über Begriffe diskutieren, man muss es sogar. Bei unklarer Theorielage sollte man sich an den differenzlogischen Ansatz des englischen Mathematikers, Logikers und Psychologen George Spencer-Brown halten, der in seinem berühmten Buch *Laws of Form* von 1969 die apodiktische Forderung »Draw a distinction« erhoben hat (SPENCER-BROWN 1997 [1969]: 3). Die Arbeit an klaren und distinkten Begriffen ist Arbeit am Verstehen der Sache. Daraus folgte der deutsche Systemtheoretiker Niklas Luhmann, dass am Beginn des differenzierten, ja, eines jeden Denkens eine Unterscheidung stehen muss. Er sagt mit Spencer Brown: »Mach eine Unterscheidung, sonst geht gar nichts. Wenn du nicht bereit bist zu unterscheiden, passiert eben gar nichts« (LUHMANN 2008: 73; vgl. dazu BÄCHLE 2016: 126). Spencer-Brown schreibt: »Unterscheidung ist perfekte Be-Inhaltung« (SPENCER-BROWN 1997 [1969]: 1). Das Nachdenken über Begriffe ist in diesem Sinn, zumindest in der Theoriebildung, ein Nachdenken über Sachverhaltsdifferenzen. Oder besser: In der Wissenschaft braucht man differenzfähige Begriffe, die sich für falsifizierbare Aussagen nutzen lassen.

Die Bildtheorie war bisher auch oft ein Tummelplatz vorwissenschaftlicher oder unsortierter Betrachtungsweisen und Begriffsverwendungen. Zu den Aufgaben der Theoriearbeit im Rahmen einer zukünftigen Bildwissenschaft gehört es daher, systematisch noch genauer auf die Schnittstellen zu modernen Theorien zu achten und bei der Begriffsverwendung viel strengere Maßstäbe einzuhalten. Oft wird auf genaue Spezifikationen verzichtet oder einfach mit dem Wort »Bild« dahinmetaphorisiert ohne Näheres zu erklären, statt genauer über Alternativen nachzudenken (die oft vorhanden sind) oder gegebenenfalls auch zu Neologismen zu greifen. Der wilde, metaphorische Gebrauch des Wortes »Bild« ist zumeist umgangssprachlich und vorwissenschaftlich motiviert, obwohl er regelmäßig auch in wissenschaftlichen oder philosophischen Publikationen stattfindet. So spricht etwa 1922 der Philosoph Ludwig Wittgenstein metaphorisch vom »logische[en] Bild«, das »die Welt abbilden« könne (*Tractatus* 2.19) oder gar vom lautsprachlichen »Satz«, der »ein Bild der Wirklichkeit« sei (*Tractatus* 4.01).

Ein Lehrstück für den beliebigen Umgang mit dem Bildbegriff in der Kunstwissenschaft bietet der Gebrauch des Wortes ›Bild‹ beim Kunsthistoriker Horst Bredekamp in seinem Freud-Buch von 2023 (BREDEKAMP 2023). Dass da von einem »Wiener Bildhauer« (89) die Rede ist, kann nicht überraschen, weil Skulpteure im Deutschen nun mal als ›Hauer des Bildes‹ bezeichnet werden. Auch der Begriff »Bildwerk« für Skulpturen ist vertretbar (35f., 84f., 129). Verwirrung hingegen stiften die uneigentlichen Wortgebräuche beim Begriff ›Bild‹. Bredekamp bezeichnet dabei immer wieder Freuds zahlreiche »Kleinskulpturen« einfach als Bild: »Einsatz der Bilder in der psychoanalytischen Praxis« in Gestalt eines »Begreifens und Betrachtens der Statuetten« (100f.); »Bildersammlung« von »Kleinskulpturen« (61); »plastisches Bildtheater« (86); »Hypnoseforschung ohne Bilder und Apparate« (120); »bildbezogene Aktivitäten« (121); die »Bilder definieren das Behandlungszimmer« (128). All das kommt dann im angeblichen »jüdischen Bilderverbot« (121f., 123) zusammen, das sich als Skulpturverbot aber eben eigentlich »auf idolhafte dreidimensionale Bilder« (121f.) bezog. Bredekamp hat auch kein Problem damit, mentale Imaginationen einfach innere »Bilder des Verdrängten« (132) zu nennen, obwohl unklar ist, was sich in der Blackbox des Bewusstseins wirklich abspielt. Wenn man diesen laxen Sprachgebrauch einem Kunsthistoriker vielleicht durchgehen lässt, so sollte eine solche terminologische Indiskretion in der Bildwissenschaft bzw. wissenschaftlichen Bildtheorie vermieden werden.

## 2.20      *Zwanzigstes Problem: Textualität*

Bildwissen und BildCode lassen sich in der Welt empirischer Beobachtungen nur indirekt erschließen; sie kommen phänomenal nicht vor. Die Quelle für ihre Rekonstruktion als abstrahierte theoretische Konstrukte im Denken der Menschen sind die empirisch vorkommenden ›Bilder‹ in konkreten kommunikativen Einbettungen. Für sie bietet sich als theoretisches Zuordnungsmodell (*genus proximum* / abstrakter Oberbegriff) im Sinne der Definitionstheorie die Kategorie ›Text‹ an (KNAPE 2024). Die *differentia specifica* von Bildtexten bezieht sich dann definitionstheoretisch gesehen allein auf die Semantik des Visuellen (vgl. WEDEWER 1985: 197–199). Eine Implikation des Axioms der Textualität ist die Kommunikabilität des Bildes.

Codes oder Sprachsysteme wirken nur auf Ebene der Textbausteine, nicht auf der Textebene selbst. Die abstrahierte Ebene des Textes muss