

## 2 Das filmische Bewegungsbild und die Temporalität der Metapher

### 2.1 Einleitung

Im ersten Teil der Studie habe ich die heuristische These der Forschergruppe *Cine-poetics* darzulegen versucht: *Eine Analyse audiovisueller Bilder muss sich auf die Interaktion zwischen Zuschauern und audiovisuellem Bewegungsbild beziehen; diese lässt sich als dynamischer, d. h. in der Zeit sich entfaltender Bildraum, beschreiben und theoretisch als filmisches Bewegungsbild fassen.*

Da wir in unseren Untersuchungen das kinematografische Bewegungsbild als generische Form aller audiovisuellen Bilder adressieren, sprechen wir von der Poiesis des *Filme-Sehens*, wenn wir uns auf diese Interaktion beziehen. Der Terminus *Filme-Sehen* lehnt sich dabei an den Alltagssprachlichen Gebrauch an, der eine kulturelle Praxis der Sinnkonstruktion qua audiovisuellen Bildern bezeichnet, die prototypisch an der Geschichte des Kinos zu fassen ist. Wir sprechen also von filmischen Bewegungsbildern (*Cinematic Movement Image*) im engen Wortsinne; und verwenden dieses Konzept mit Blick auf alle möglichen audiovisuellen Bilder – über die eigentliche Domäne des Kino- oder Dokumentarfilms hinaus –, sofern diese über die unmittelbare audiovisuelle Reproduktion hinaus ein Mindestmaß an rhetorischer respektive poetischer Gestaltung aufweisen.

Im zweiten Teil möchte ich nun versuchen, die Konsequenzen eines solchen Ansatzes für die methodologischen Modelle kognitiver Verarbeitungsschemata oder bedeutungsgenerierender Prozesse auszuarbeiten, die regelmäßig in Anschlag gebracht werden, wenn die Diskurse und Sinnkonstruktionen audiovisueller Bewegungsbilder untersucht werden. Wenn ich mich dabei insbesondere auf die kognitionstheoretischen Modelle beziehe, dann nicht nur, weil sie die methodologischen Debatten zur Frage des Verstehens und der Bedeutungsgenerierung audiovisueller Bilder seit geraumer Zeit dominieren. Sie bezeichnen auch den strategischen Ort, an dem der Diskurs audiovisueller Bilder von allen Aspekten kultureller Produktion gereinigt wird, um ihn als Strom von Informationsprozessen für die wissenschaftliche Untersuchung beherrschbar zu machen.<sup>1</sup> Mir geht

---

<sup>1</sup> Vgl. de Certeau Begriff der Strategie im Unterschied zur Taktik: Certeau: Kunst des Handelns, S. 87 f. Vgl. zudem: „Wie beim Management ist jede strategische Rationalisierung vor allem darauf gerichtet, das Umfeld von dem eigenen Bereich, das heißt vom Ort der eigenen Macht und des eigenen Willens, abzugrenzen. [...] Ein Unterfangen der Moderne in politischer, militärischer und wissenschaftlicher Hinsicht.“ Certeau, S. 88.

es dabei keineswegs um eine kulturwissenschaftliche Kritik, die meint, sich mit dem Hinweis auf ihren konstruktivistischen Fundamentalismus aus der Diskussion verabschieden zu können. Mein Ziel ist vielmehr – ein sicherlich vermessenes Unterfangen – die *taktische Durchquerung* dieses Ortes, um in der Schwebelage von Kritik und Aneignung der dort versammelten Argumentationsstrategien und Beschreibungsmodelle den Blick für eine kulturwissenschaftliche Theorie des Denkens filmischer Bewegungsbilder zu öffnen; in der Hoffnung, damit die Grundlagen zu erschließen, die eine methodische Untersuchung der Geschichte der Poiesis des Filme-Sehens, d. i. der Kunst des Konsums audiovisueller Bewegungsbilder, ermöglichen.

Die Poiesis des Filme-Sehens lässt sich nämlich nicht – das wurde bereits weiter oben erläutert – im Rahmen von Kommunikationsmodellen beschreiben, die von fixen Akteurspositionen ausgehen (Produzenten, Rezipienten); diese reproduzieren in immer neuen Formeln die Aporien des traditionellen Sender-Empfänger-Modells, der Triade von Codierungssystem, Codierung und Dekodierung. Ist doch sowohl auf Seiten der Produktion wie auf Seiten der Rezeption von einer Pluralität auszugehen, die sich letztlich nur als Netzwerk zwischen Medien, Akteuren, Bedeutungssystemen und kulturellen Sinnbeständen sinnvoll beschreiben lässt. Entsprechend meint die Poiesis des Filme-Sehens *die Interaktion einer unbestimmbaren Pluralität von Rezipienten mit audiovisuellen Bildern einer ebenso unbestimmbaren Pluralität kultureller Produktion, die wiederum durch eine Vielzahl medialer Operationen und Funktionen verbunden sind.*<sup>2</sup>

Damit verlagert sich die Frage der Sinnkonstruktion weg von der Ebene (ab-) bildlicher Repräsentation – sei es der Dingwelt, sei es der Folge von Handlungsszenarien – auf die der Interaktion von Rezipient und audiovisuellem Bewegungsbild. Die Interaktion zwischen Bewegungsbild und Wahrnehmungsprozessen leibhafter Zuschauer ist als temporale Struktur filmischer Bewegungsbilder analytisch zu fassen und als *innere* Dynamik des in der Zeit sich entfaltenden Bildraums zu beschreiben. Damit aber ist der Wirklichkeitsbezug des Bildes selbst noch der Dynamik dieser Interaktion unterworfen. D. h., er ist zur Funktion der Bildraumkonstruktion geworden. Zuallererst gilt es, den Prozess der Interaktion als dynamische Entfaltung eines Bildraums analytisch zu rekonstruieren, der dann seinerseits in seinem spezifischen Wirklichkeitsbezug bestimmt werden muss.

---

<sup>2</sup> Vgl. Schüttpelz/Thielmann (Hrsg.): Akteur-Medien-Theorie.

## Fingierte Wahrnehmungsszenarien

Was aber bedeutet es, wenn man mit Blick auf die Analyse audiovisueller Bilder weder von einer selbstverständlich gegebenen Alltagswirklichkeit, noch von universellen kognitiven Schemata ausgeht? Wie lässt sich die Poiesis des Filmsehens überhaupt erfassen, wenn sie nicht in der Abfolge isolierter audiovisueller Bild-Repräsentationen zu bestimmen ist? Lassen sich über einen Rezeptionsprozess überhaupt methodisch distinkte Aussagen treffen, wenn dieser als eine Interaktion von Bild und Rezipient modelliert ist, die in ihren operativen Vernetzungen zwischen Affekten, kognitiven Schemata, kulturellen Semantiken und medialen Operationen nach allen Seiten hin offen ist?

Um diese Fragen beantworten zu können, ist die grundlegende Prämisse der kognitionstheoretischen Ansätze zum Film-Verstehen in einem entscheidenden Aspekt zu korrigieren. Wenn diese Prämisse nämlich lautet, dass Zuschauer im Prozess des Film-Sehens eine diegetische Welt konstruieren, zu der sie sich verhalten, *als ob* sie mit einer real sie umgebenden Welt interagierten,<sup>3</sup> *bezeichnet genau dieses ‚Als-ob‘ den kritischen Punkt; es weist auf einen Prozess des Fingierens zurück, auf das Herstellen einer Fiktion;*<sup>4</sup> der Prozess des Fingierens aber ist weder im Modell der Kommunikation von Bedeutungen und Narrativen, noch in der Repräsentation von dem Bild vorgängigen Sachverhalten zu fassen.

Die Fiktion des ‚Als-ob-ich-mit-einer-real-mich-umgebenden-Wirklichkeit-interagierte‘ ist eben als Interaktion von Affekten, kognitiven Schemata und kulturellen Semantiken zu rekonstruieren, die sich im Wechselspiel von audiovisuellem Bewegtbild und empfindsamen Körpern der Rezeption vollzieht. Im Fingieren erschließt sich die Audiovisualität des Bewegtbildes als ein Feld unbestimmter bzw. unterbestimmter Sinneseindrücke, die zum Einfallstor ‚taktischer Feinheiten‘ im Sinne de Certeaus werden; den Rezipienten wird das Bild zu einem Wahrneh-

---

3 Dieser Aspekt filmischer Rezeption steht in der Diskussion um die Immersion im Zentrum. Vgl. Marie-Laure Ryan: *Narrative as Virtual Reality. Immersion and Interactivity in Literature and Electronic Media*, Baltimore 2001; Oliver Grau: *Virtual Art. From Illusion to Immersion*, Cambridge, Mass. 2003; Alison Griffiths: *Shivers Down Your Spine. Cinema, Museums, and the Immersive View*, New York 2008; Robin Curtis: *Immersion und Einfühlung. Zwischen Repräsentationalität und Materialität bewegter Bilder*, in: *montage/av* 17 (2), 2008, S. 89–107; Christiane Voss: *Fiktionale Immersion*, in: „Es ist, als ob“. *Fiktionalität in Philosophie, Film- und Medienwissenschaft*, hrsg. v. Gertrud Koch, Christiane Voss, Paderborn 2009, S. 127–138; Valentijn T. Visch/Ed S. Tan/Dylan Molenaar: *The Emotional and Cognitive Effect of Immersion in Film Viewing*, in: *Cognition and Emotion* 24 (8), 2010, S. 1439–1445.

4 Gertrud Koch/Christiane Voss (Hrsg.): „Es ist, als ob“. *Fiktionalität in Philosophie, Film- und Medienwissenschaft*, Paderborn 2009; Koch: *Die Wiederkehr der Illusion*, S. 77.

mungsszenario, in das sie sich einnisten, das sie in ihrer Sinnestätigkeit, ihren Begierden, Affekten und Gedanken besetzen können – wie ein fremdes Haus, das man besetzt und für seine eigenen Zwecke umräumt, um es zu bewohnen, um sich selbst in der Inbesitznahme zu genießen.<sup>5</sup> Die Bildkonsumenten führen in der taktischen Aneignung ihre eigene Subjektivität – das Genießen ihrer Subjektivität – als Differenz in die Schemata und Affektskripte der Alltagswahrnehmung ein. Diese Differenzierungen betreffen sehr wohl die Strukturen kognitiver Verarbeitung der audiovisuellen Bewegtbilder in ihrer je spezifischen medienästhetischen Ausprägung. Darf man doch annehmen, dass das ästhetische Vergnügen die Prozesse des Sehens und Hörens bis in die Abläufe der Sinnesverarbeitung verändert. (Ein Indiz dafür mag gerade in der allgegenwärtigen Diagnose der Pornosucht zu finden sein; einer Sucht also, die sich auf audiovisuelle Bilder als einen Stoff bezieht, der unmittelbar sexuell erregend wirkt. Scheint hier doch offensichtlich ein Reiz-Reaktionsschema zwischen technischem und humanem Körper wirksam zu sein, das tendenziell die Differenzen und Zwischenräume, die im fingierenden Als-ob entstehen, auszulöschen vermag. Bilder, deren wirkungsästhetisches Kalkül den Konsumenten tendenziell den Spielraum taktischer Aneignung zu entziehen sucht.)

Ich schlage deshalb vor, den Prozess der Sinnbildung audiovisueller Bilder nicht als handlungslogische Rekonstruktion repräsentierter Ereignisfolgen, sondern als ein Fingieren, ein Herstellen fiktiver Wahrnehmungswelten im Prozess der Rezeption zu verstehen. Die Verarbeitung audiovisueller Daten im Modus des *Als-ob* lässt eine Welt entstehen, deren Sinneslogik und Sinnhaftigkeit nicht per se gegeben ist, sondern von Zuschauern, der Axiomatik je spezifischer Bildraumkonstruktionen folgend, als eine Art und Weise des Sehens, Hörens, Fühlens und Denkens zu generieren ist. Die Übersetzung der Sinnesdaten audiovisueller Bewegtbilder in die konkrete Sinnlichkeit fingierter Wahrnehmungswelten – die unserer Alltagswelt gleichen, ohne mit ihr identisch zu sein – ist der basale Prozess, in dem jedes Verstehen audiovisueller Bewegtbilder seinen Ausgang nimmt. Ein solcher Aneignungsprozess, eben die *Poiesis des Filme-Sehens*, lässt sich als eine Ausdifferenzierung der kognitiven Schemata unserer Alltagswahrnehmung in immer neuen Raum-Zeit-Figurationen<sup>6</sup> verstehen. Darauf bezieht sich die im ersten Teil des Buches erläuterte Rede vom ‚Denken filmischer Bilder‘.

<sup>5</sup> Siehe de Certeaus Vergleich mit einer Mietwohnung: Certeau: *Kunst des Handelns*, S. 27.

<sup>6</sup> Deleuze: *Das Bewegungs-Bild*, S. 13–26; Deleuze: *Das Zeit-Bild*, S. 95–167; Massumi: *Parables for the Virtual*, S. 23–88; Brian Massumi: *Ontomacht. Kunst, Affekt und das Ereignis des Politischen*, Berlin 2010, S. 69–104.

Man erinnere sich noch einmal der Exposition von ČELOVEK S KINOAPPARATOM; jenem Film, der wie wenige andere für die Erforschung der Möglichkeiten des filmischen Bewegungsbildes durch die ästhetischen Avantgarden des vergangenen Jahrhunderts eintreten kann. Wenn Vertov das schlafende Mädchen im Federkissen inszeniert, dann lassen sich die einzelnen Einstellungen in der Montagefiguration des Bewegungsbildes auf vergleichbare Weise als stetig fortschreitende Differenzierungen eines basalen Schemas beschreiben, die sich wie die Worte ‚Bett‘, ‚Lager‘ und ‚Ruhestätte‘ zum perzeptiven Schema ‚Schlafplatz‘ verhalten. In der Poiesis des Filme-Sehens entsteht die Fiktion, man sehe einer Stadt bei ihrem Erwachen zu, ja man empfinde sich selbst als ein Körper, der Anteil habe an diesem Erwachen. Der Prozess dieser Fiktionalisierung setzt keineswegs mit der per se verständlichen Bild-Repräsentation ein. Er bezieht sich vielmehr auf die temporale Struktur des Bewegungsbildes selbst, in der in diskrete Perzepte zerlegte Schemata der Alltagswahrnehmung, der Bildgestalt und der symbolischen Form miteinander verschränkt und aufeinander bezogen sind. Der diegetische Horizont ist keineswegs im linearen Zeitschema kausallogischer Verknüpfung apriorisch vorauszusetzen, sondern muss als immanente Logik der Verknüpfung höchst unterschiedlicher dynamischer Raum-Zeit-Figurationen vom Zuschauer erfasst und zugleich als Einheit einer fiktionalen filmischen Welt entworfen werden.

Die Bild-Repräsentation ‚eine Frau, die schläft und langsam erwacht‘ ist das Ergebnis eines Prozesses der Fiktionalisierung, der durch alle Register medientechnischer und ästhetischer Modulationen hindurchgegangen ist. Das filmische Szenario ‚eine Frau erwacht‘ ist eine Bewegungsfiguration, eine rhythmische Figur aus Lichtwerten und grafischen Mustern – mindestens ebenso sehr wie eine Szene, in der eine Frau in ihrem Bett schläft und langsam erwacht.

An ČELOVEK S KINOAPPARATOM lässt sich anschaulich nachvollziehen, was im Prinzip für alle Filme gilt, seien sie der dokumentarischen, experimentellen oder der unterhaltenden Sparte zuzuordnen. Sie alle setzen voraus, dass die Sinneslogik und Sinnhaftigkeit der wahrgenommenen Welt im Filme-Sehen zuallererst hervorgebracht, als Wahrnehmungswelt fingiert werden muss, bevor sie gemäß referentieller Relationen zwischen der filmischen Bildraumkonstruktion und unserer Alltagswelt als dokumentarisch, fiktional oder experimentell klassifiziert werden kann.<sup>7</sup>

---

<sup>7</sup> In diesem Punkt scheint sich die literarische Narration nicht von der filmischen zu unterscheiden. Vgl. Armen Avanesian/Anke Hennig: *Präsens. Poetik eines Tempus*, Zürich 2012; Armen Avanesian/Anke Hennig: *Metanoia. Spekulative Ontologie der Sprache*, Berlin 2014. Wenn Ge-

Daraus ergibt sich eine weitere heuristische These: *Wenn Zuschauer sich zur im Film dargestellten Welt in bestimmten affektiven und perzeptiven Aspekten verhalten, als ob die filmische Welt eine sie umgebende Alltagswelt wäre, geschieht dies immer schon in der reflexiven Wendung auf einen vorgängigen Prozess der Fiktionalisierung des Bildes selbst.* Der Prozess der Fiktionalisierung setzt, im Durchgang durch alle Register medientechnischer und ästhetischer Modulationen, die Schemata der Verarbeitung symbolischer Formen (Sprache, darstellendes Spiel, Ikonografien, rhythmische Bewegungsfigurationen, Montageeinheiten) in Beziehung zu Schemata von Bewegung, Aktion und Perzeption, die in der Alltagswahrnehmung gründen, ohne dass sich das eine (die symbolischen Formen) auf das andere (die kognitiven Schemata) zurückführen ließe.

Erst in der fiktionalisierenden Extension der Poiesis des Filme-Sehens werden die audiovisuellen Bilder zu filmischen Bewegungsbildern raum-zeitlicher Konfigurationen. Erst in diesen Konfigurationen fügt sich die Folge audiovisueller Bilder zu einer diegetischen Welt, die sich als Repräsentation verstehen lässt. Die dargestellte filmische Welt ist nicht Ausgangs-, sondern Endpunkt des Fiktionalisierungsprozesses. Sie weist nicht auf Bild-Repräsentationen zurück, sondern geht aus der Poiesis des Filme-Sehens als ein Denken in veränderlichen Raum-Zeit-Figuren hervor.

### Die Metapher als Königsweg

Wie aber lässt sich ein solcher Fiktionalisierungsprozess modellieren? In unserer Arbeit am Forschungskolleg *Cinepoetics* haben wir eine Antwort auf diese Frage in der Auseinandersetzung mit den Theorien zur Metapher gesucht. Kommt doch mit dem fingierenden *Als-ob* ein Erfahrungsmodus in den Blick, der auch zu den grundlegenden Bestimmungen der Metapher gehört. Die Metapher nämlich – um eine der prominentesten Positionen der Metaphertheorie aufzurufen – verbindet zwei Erfahrungsbereiche durch eine solche *Als-ob*-Relation: Sie lässt sich fassen als „understanding and experiencing one kind of thing in terms of another“.<sup>8</sup> Dieser Prämisse folgend, wird die metaphorische Emergenz von Bedeutungen selbst als ein Phänomen der Übertragung verkörperter Aktions- und Bewegungsschemata zwischen unterschiedlichen Erfahrungsbereichen rekonstruiert.<sup>9</sup>

---

nette davon spricht, dass die Narration eine Extension des Verbs darstellt, impliziert auch dies einen Prozess der Fiktionalisierung; vgl. Gérard Genette: *Die Erzählung*, München 1998, S. 18.

<sup>8</sup> George Lakoff/Mark Johnson: *Metaphors We Live By*, Chicago 1980, S. 3.

<sup>9</sup> Raymond W. Gibbs: *Embodiment and Cognitive Science*, Cambridge 2006.

Übertragen auf die Metaphorizität audiovisueller Bilder vollziehen in dieser Perspektive die Zuschauer tatsächlich selbst den Vorgang fikionalisierenden Dichtens; sie lassen in den Modi der Metaphorisierung, d. h. dem „seeing as“<sup>10</sup>, „seeing through“ oder „seeing one thing in terms of another“, die isolierten audiovisuellen Bild-Repräsentationen zu filmischen Bewegungsbildern werden, die in jedem einzelnen Zug ihrer Ausdrucksdimension und ihrer Performanz eine künstlich gemachte, eine hergestellte, eine fiktive Wirklichkeit entstehen lassen. In den Modi des „seeing as“<sup>11</sup>, „seeing through“ oder „seeing one thing in terms of another“ wird ein Bildraum entworfen, ein Rahmen raum-zeitlicher Relationen gesetzt, es werden die Positionen festgelegt, innerhalb derer sich das Verstehen als Rekonstruktion einer Handlung, einer Erzählung, einer Dokumentation erst entfalten kann. Das ist der Grund, weshalb die Rezeption filmischer Bilder als eine Poiesis im engen Sinne begriffen werden kann: als ein Akt dichterischer Produktion, der im medialen Konsum selber zu verorten ist.<sup>12</sup> Deren grundlegender Modus ist die metaphorische Übertragung auf der Basis körperlicher Wahrnehmungsschemata und -prozesse. Den Übertragungsprozess selbst hat einer der wichtigsten Vertreter moderner Metapherntheorie, nämlich Max Black, als Interaktion konzipiert. Seinen Einsichten folgend, kann *die Metapher den Königsweg zum Verständnis der Sinnkonstruktionen audiovisueller Bilder* weisen.<sup>13</sup>

Wie nun gelangt man von audiovisuellen Sinnesdaten zu filmischen Bildern? Wie kommt man von der Interaktion mit audiovisuellen Bildern zu den Sinnkonstruktionen komplexer Bildräume? Auch wenn die Geschichte der Filmtheorie von Eisenstein bis Deleuze sich zwanglos als Genese einer Poetologie, d. i. einer Theorie der Poiesis des Filme-Sehens, lesen lässt, steht eine Methodologie im strengen Sinne aus.

In diesem Punkt hat die Theorie der Metapher – oder besser: der *metaphorischen Konzepte* (das ist der Terminus, den Lakoff und Johnson ursprünglich verwenden)<sup>14</sup> – der Theorie audiovisueller Bilder etwas voraus. Hat doch die *Conceptual Metaphor Theory (CMT)* zu einer breiten methodologischen Diskussion und entsprechenden Analysen und Analysemodellen geführt. Mit ihr im Rücken

<sup>10</sup> Vgl. zum ‚seeing as‘ die Wittgensteinsche Duck-Rabbit-Figur. Ludwig Wittgenstein: Philosophische Untersuchungen, in: ders.: Schriften Band 1. Tractatus logico-philosophicus. Tagebücher 1914–1916. Philosophische Untersuchungen, Frankfurt/M. 1969, S. 279–544, hier S. 504.

<sup>11</sup> Wittgenstein, S. 504.

<sup>12</sup> Certeau: Kunst des Handelns.

<sup>13</sup> Vgl. den schönen Aufsatz von Hartmut Winkler: Metapher, Kontext, Diskurs, System, in: Kodikas/Code. Ars Semeiotika 12 (1–2), 1989, S. 21–40, online abrufbar unter <http://homepages.uni-paderborn.de/winkler/metapher.html> (29. 03. 18).

<sup>14</sup> Lakoff/Johnson: Metaphors We Live By.

konnte sich die angewandte Metaphernforschung als eigenes Forschungsgebiet international etablieren.<sup>15</sup>

In gewisser Weise erging es den Metaphern lange Zeit ähnlich wie den audiovisuellen Bildern. Sie waren als ‚Bild‘ oder bildhafter Ausdruck allzu selbstverständlich zu verstehen. Und wenn sie denn doch Gegenstand theoretischer Überlegungen wurden, geschah dies häufig im taxonomischen Vergleich mit anderen sprachlichen Ausdrucksformen. Auch wenn mit Nietzsche die Metapher zum Paradigma eines radikal rhetorisch gewendeten Wahrheitskonzepts wurde<sup>16</sup> – als Denkform eigenen Rechts trat sie erst in der zweiten Hälfte des zwanzigsten Jahrhunderts in Erscheinung. Die Auseinandersetzung mit der entsprechenden Diskussion mag Hinweise liefern, um einen analogen Schritt bezogen auf das audiovisuelle Bild zu initiieren, mit dem Ziel, eine vergleichbare methodologische Kehre hinsichtlich der Analyse (kultureller) Sinnbildungsprozesse zu vollziehen, die durch die Interaktion mit audiovisuellen Bildern geprägt sind.

### Konzeptuelle Metaphertheorie

Wenn sich in den letzten Jahren die Forschung im Feld der *CMT* verstärkt den audiovisuellen Bildern zuwendet, dann geschieht dies nicht nur eingedenk der offensichtlichen gesellschaftlichen Relevanz audiovisueller Medien für die politischen, sozialen, ökonomischen, wissenschaftlichen, religiösen und pädagogischen Öffentlichkeiten und Diskurse. Das darin zum Ausdruck kommende Erkenntnisinteresse ist durchaus intrinsisch motiviert. Ist doch die These, dass Metaphern kognitive Konzepte vor aller sprachlichen Ausdruckform sind, das *Movens* der theoretischen Debatten im Umkreis der *CMT*. Und diese These findet sich nicht zuletzt in den Untersuchungen zur *Multimodalität* menschlicher Ausdrucksformen (hier vor allem der Geste) mit größter Evidenz bestätigt. Denn sowohl mit Blick auf körperliche Gesten wie auf audiovisuelle Bewegungsbilder lassen sich Prozesse metaphorischer Bedeutungsgenerierung aufweisen, die

---

<sup>15</sup> Vgl. die Geschichte der RaAM (Association for Researching and Applying Metaphor): Raymond W. Gibbs: The Wonderful, Chaotic, Creative, Heroic, Challenging World of Researching and Applying Metaphor: A Celebration of the Past and some Peeks into the Future, in: Researching and Applying Metaphor in the Real World, hrsg. v. Graham Low, Zazie Todd, Alice Deignan, Lynne Cameron, Amsterdam/Philadelphia 2010, S. 1–18.

<sup>16</sup> Vgl. Friedrich Nietzsche: Über Wahrheit und Lüge im außermoralischen Sinne, in: ders.: KSA I. Die Geburt der Tragödie, Unzeitgemäße Betrachtungen I–IV, Nachgelassene Schriften 1870–1873, hrsg. v. Giorgio Colli und Mazzino Montinari, München 1988, S. 876–890.



nicht mehr in sprachlichen Ausdrucksformen formuliert sind.<sup>17</sup> Die *CMT* – das hat die Gestenforschung gezeigt<sup>18</sup> – eröffnet Möglichkeiten, Prozesse der Bedeutungsgenerierung zu beobachten, die primär durch körperliche Interaktion, affektive Verwicklung und unbewusste Synchronisierungen gestischen Ausdrucksverhaltens strukturiert sind.

Nun umfasst die *CMT* in all ihren Verästelungen nur einen spezifischen Bereich der aktuellen wissenschaftlichen Forschung zur Metapher. Tatsächlich ist das Feld weit gefächert und schließt höchst disparate Positionen in sich ein. Grob vereinfacht kann man aber sagen, dass es zur einen Seite durch die kogniti-

---

**17** Die Analyse eines Gruppengesprächs, das zum Thema Selbstverwirklichung geführt wurde, zeigt, wie ein metaphorisches Szenario in der Interaktion der drei Gesprächsteilnehmerinnen entsteht: Im Zusammenspiel von Gestik und Sprache entsteht ein geteilter inter-affektiver Parcours, den die Gesprächsteilnehmerinnen durchlaufen. In einer Phase des Gesprächs, in der sich eine der Diskutantinnen unverstanden fühlt, da die anderen beiden ihre Idee von Selbstverwirklichung als Form des Zwangs ablehnen, untermalt sie die Äußerung, dass hinter der Selbstverwirklichung nur ein ‚Tu-dies-und-tu-das‘ stünde, mit den Lauten „Hopp, Hopp“ und einer peitschenden Handbewegung: „The ‚explosion‘ of the ‚hopp hopp‘ gestures (and the exclamation) is just one facet of a metaphorical scenario that structures the discussion around the idea of a metaphorical whipping activity as positive incentive for engaging in activities relevant for one’s own self-realization or as destructive intrusion into a free self.“ (Cornelia Müller/Hermann Kappelhoff: *Cinematic Metaphor. Experience – Affectivity – Temporality*, Berlin/Boston 2018 (in Vorbereitung), Kapitel 9). Vgl. auch Dorothea Horst/Franziska Boll/Christina Schmitt/Cornelia Müller: *Gesture as Interactive Expressive Movement: Inter-Affectivity in Face-to-Face Communication*, in: *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 2, hrsg. v. Cornelia Müller, Alan Cienki, Ellen Fricke, Silva H. Ladewig, David McNeill, Jana Bressemer, Berlin/Boston 2014, S. 2112–2125.

**18** Vgl. Cornelia Müller: *Waking Metaphors. Embodied Cognition in Multimodal Discourse*, in: *Metaphor. Embodied Cognition in Discourse*, hrsg. v. Beate Hampe, Cambridge 2017, S. 297–316; Horst/Boll/Schmitt/Müller: *Gesture as Interactive Expressive Movement*; Cornelia Müller/Silva H. Ladewig: *Metaphors for Sensorimotor Experiences: Gestures as Embodied and Dynamic Conceptualizations of Balance in Dance Lessons*, in: *Language and the Creative Mind*, hrsg. v. Michael Borkent, Barbara Dancygier, Jennifer Hinnell, Stanford 2013, S. 295–324; Astrid Kolter/Silva H. Ladewig/Michela Summa/Cornelia Müller/Sabine C. Koch/Thomas Fuchs: *Body Memory and the Emergence of Metaphor in Movement and Speech. An Interdisciplinary Case Study*, in: *Body Memory, Metaphor, and Movement*, hrsg. v. Sabine Koch, Thomas Fuchs, Cornelia Müller, Amsterdam/Philadelphia 2012, S. 201–226; Cornelia Müller: *Gesten als Lebenszeichen ‚toter Metaphern‘*, in: *Zeitschrift für Semiotik* 25 (1–2), 2003, S. 61–72; Eve Sweetser: *Regular Metaphoricity in Gesture: Bodily-based Models of Speech Interaction*, Conference Paper, Actes du 16e Congrès International des Linguistes, Elsevier 1998; Alan Cienki: *Conceptual Metaphor Theory in Light of Research on Speakers’ Gestures*, in: *Cognitive Semiotics* 5 (1–2), 2009, S. 349–366; Irene Mittelberg/Gina Joue: *Source Actions Ground Metaphor via Metonymy: Toward a Frame-based Account of Gestural Action in Multimodal Discourse*, in: *Metaphor. Embodied Cognition and Discourse*, hrsg. v. Beate Hampe, Cambridge 2017, S. 119–137.

onpsychologische und neurowissenschaftliche Forschung zur Theorie kognitiver Konzepte und Schemata begrenzt ist; zur anderen Seite durch die philosophische Anthropologie, die sich auf die Geschichte als Geschichte der kulturellen oder geistesgeschichtlichen Entstehung der Formen menschlichen Denkens bezieht. Zugespitzt kann man also von zwei repräsentativen Ansätzen sprechen, die sich diametral in ihrem Erkenntnisanspruch entgegenstehen. Denn beide nehmen für sich in Anspruch, die Formen menschlichen Denkens zu erforschen – einmal als geschichtlich rekonstruierbare Entwicklung, einmal als naturwissenschaftlich zu beschreibendes kognitives System.

Auf der einen Seite steht die *CMT*, die aktuell vor allem im Feld der Linguistik und der Kognitionswissenschaften den Diskurs zur Metapher dominiert; dieser theoretische Ansatz definiert sich explizit über einen radikalen Bruch mit Rhetorik und Poetik menschlicher Ausdrucksformen.<sup>19</sup> Auf der anderen Seite steht die *Metaphorologie*, welche gerade versucht, in der Metapher die rhetorische Dimension sprachlicher Äußerungen selbst in ihrer erkenntnistheoretischen Relevanz zu fassen. Ihre Grundannahme lautet, dass wir die Welt entsprechend der medialen und symbolischen Instrumente ihrer Repräsentation verstehen.

Bezogen auf die im ersten Teil skizzierte Theorie audiovisueller Bilder scheinen beide Perspektiven vorderhand so unverzichtbar, wie sie unvereinbar scheinen. Bringt doch zum einen die *CMT* ein Modell verkörperter Erfahrung ins Spiel, das es ermöglicht, Verstehensprozesse in ihrem Rekurs auf konkrete sinnliche, d. h. verkörperte Wahrnehmungs- und Aktionskonstellationen zu beschreiben. Sie spricht damit unmittelbar den konstitutiven Aspekt des Verstehens audiovisueller Bilder an: Denn – so habe ich im ersten Teil zu zeigen versucht – eben die Verortung der Sinnkonstruktion eines sich selbst als gegenwärtig erfahrenden sinnlich-konkreten Wahrnehmungserlebens grundiert alle Prozesse filmischen Verstehens und Denkens. Demgegenüber verspricht die *Metaphorologie* einen Zugang zur Geschichte des Entstehens der medialen, ästhetischen und sprachlichen Formen menschlichen Denkens; sie eröffnet den Zugang zu einem Verständnis von Wirklichkeit, demzufolge diese immer die geteilte Wirklichkeit einer kulturellen Gemeinschaft ist. Jede Theorie audiovisueller Bilder, die Anspruch darauf erhebt, die Bilder selbst als genuine Formen menschlichen Denkens zu begreifen und deren Funktion für die Modellierung kultureller Sinnhorizonte zu erschließen, ist innerhalb einer solchen historischen Perspektive zu situieren.

Ist doch – so lautet die für meine Überlegungen zentrale Ausgangsthese – aus kulturwissenschaftlicher Perspektive *jegliche Frage zu Metaphern und meta-*

---

<sup>19</sup> Vgl. Charles Forceville/Eduardo Urios-Aparisi (Hrsg.): *Multimodal Metaphor*, Berlin/Boston 2009; Forceville: *Non-Verbal and Multimodal Metaphor in a Cognitivist Framework*.

*phorischen Konzepten in audiovisuellen Bildern untrennbar mit der Frage nach deren Historizität, d. h. ihrer kulturhistorischen Verortung, verbunden.* Vor diesem Hintergrund kann es kein Verständnis der Sinnbildungsprozesse audiovisueller Bilder geben, welches die grundsätzliche Geschichtlichkeit kultureller Sinnbildung ausblendet.<sup>20</sup> Bereits eine systematische Analyse, die lediglich darauf abzielt, die Interaktion von Zuschauerinnen und Zuschauern mit audiovisuellen Bildern als einen Prozess des *meaning-making* im Sinne der Sprachwissenschaft zu rekonstruieren – d. h., den Dynamiken der Versprachlichung audiovisueller Bilder nachzugehen –, wird die historische Dimension einer solchen Analyse stets berücksichtigen müssen. Warum dem so ist, wird noch genauer zu bedenken sein.

Zuvor wäre jedoch zu präzisieren, worin denn das Erkenntnisinteresse besteht, das die medienwissenschaftliche Forschung an den gegenwärtigen Diskussionen zur Theorie der Metapher hat. Was kann die Medienwissenschaft von der Metaphernforschung lernen? Und umgekehrt, wie kann die weitläufige Forschung zu Metaphern in Bildern audiovisueller Medien von der Theorie filmischer Bilder profitieren? Lassen sich die unterschiedlichen Konzepte von Metaphern überhaupt aufeinander beziehen? Oder haben sie vielleicht umgekehrt nicht viel mehr als die Benennung eines Phänomens gemeinsam, welches in der Folge als theoretischer Gegenstand auf völlig verschiedene und inkompatible Weise definiert wird?

Um diese Frage zu beantworten, möchte ich mich im Folgenden zunächst und vor allem mit der *CMT* auseinandersetzen. Trotz oder gerade weil für das eigene Verständnis audiovisueller Bilder die Historizität ein konstitutives Axiom jeder Sinnkonstruktion darstellt. Erst im Anschluss daran werde ich die Eckpunkte einer Theorie der Bewegungsbild-Metaphern darlegen, die auf eine historische Poetologie audiovisueller Bilder abzielt. Auch wenn ich mich dabei auf einen abschließenden Ausblick beschränke, handelt es sich doch um den Fluchtpunkt aller folgenden Argumente.

## 2.2 Zur Kritik der Anwendung der *CMT*

Auf den ersten Blick entspricht der Gegensatz von kognitiver Metaphertheorie und Metaphorologie einer Konstellation, die sich analog zu dem im ersten Teil des Buches behandelten Verhältnis von kognitiver Filmtheorie und der (medien-)his-

---

<sup>20</sup> Zum Verhältnis von Film und Geschichtlichkeit vgl. die Arbeiten Michael Wedels. Wedel: Der deutsche Musikfilm; Michael Wedel: Filmgeschichte als Krisengeschichte. Schnitte und Spuren durch den deutschen Film, Bielefeld 2011.

torischen Poetologie audiovisueller Bilder verhält. Nur dass im Fall der *CMT* die Verbindung zur Poetik offensichtlich gekappt wird, während der Neoformalismus diese explizit zu wahren suchte.

So führt etwa Charles Forceville, einer der führenden Vertreter der analytischen Anwendung der *CMT* auf audiovisuelle Medien,<sup>21</sup> aus:

From [Lakoff and Johnson's *CMT*] on, metaphor was no longer one of a series of tropes that could enhance or embellish the aesthetic meaning of a poem or the persuasive power of speeches (as had been its primary claim to fame since Aristotelian times), but one of the conceptual tools for human beings to make sense of the world.<sup>22</sup>

Forcevilles Einordnung zielt primär auf eine theoriegeschichtliche Verortung der *CMT* als Ursprung der Metaphernforschung als eigenständiger Disziplin.<sup>23</sup> Der Metapher wird gegenüber anderen sprachlichen Tropen und rhetorischen Formen eine Sonderstellung eingeräumt; sie wird als paradigmatisches Muster kognitiver Prozesse verstanden. Freilich wird mit der ausschließenden Entgegensetzung der Metapher als eines kognitiven Schematismus und der Metapher in Rhetorik und Poetik eine massive Eingrenzung des Gegenstandsfeldes vorgenommen. Durch den Bruch mit den theoriegeschichtlich prägenden Traditionen der Poetik und Rhetorik werden alle poetischen und rhetorischen Aspekte der metaphorischen Konstruktion aus dem Fokus entsprechender Metaphernanalysen gerückt.

Dies bringt zunächst einmal eine willkommene Vereinfachung der analytischen Perspektive mit sich.

### **Probleme der Anwendung der *CMT* auf audiovisuelle Medien**

Die strikte Abgrenzung von Rhetorik und Poetik ermöglicht es, universelle kognitive Analyseschemata – unbesehen ihrer tatsächlichen kulturellen Formung – auf alle Phänomene menschlicher Kommunikation anzuwenden, ohne sich mit den Differenzen zwischen symbolischen Modi (seien es Bild-, Sprach-, oder Bewegungsbildformen), noch viel weniger mit deren medialer, kultureller und historisch spezifischer Ausprägung auseinandersetzen zu müssen. Das aber hat massive Konsequenzen für die Analyse audiovisueller Bilder.

---

**21** Charles Forceville: Visual and Multimodal Metaphor in Film: Charting the Field, in: *Embodied Metaphors in Film, Television, and Video Games. Cognitive Approaches*, hrsg. v. Kathrin Fahlenbrach, New York 2016, S. 17–32.

**22** Forceville, S. 17.

**23** Forceville, S. 17–18.

Rückt man die ästhetische und die – für die Rhetorik zentrale – persuasive Dimension audiovisueller Bilder als akzidentielle Zutat beiseite, gerät der Umstand, dass audiovisuelle Bewegtbilder zuerst und vor allem wahrgenommen und als Wahrnehmungsereignis von konkreten Rezipienten verkörpert sein müssen, bevor sie überhaupt verstanden werden können, völlig aus dem Blick. Gerade der für die CMT zentrale Aspekt des *embodied meaning* wird zur leeren Formel, wenn man von dem Wahrnehmungsereignis absieht, das Bewegtbilder zuerst und vor allem initiieren und stattdessen die Bild-Repräsentationen zum Ausgangspunkt der Metaphernanalyse macht – als fügten diese sich wie Worte und Sätze zu einem sprachlichen Text, der für jedermann lesbar ist.

Folgerichtig ergeben sich aus der im ersten Teil entwickelten bildtheoretischen Perspektive zunächst fundamentale Einwände gegen die CMT-gestützte Analyse audiovisueller Metaphern. Jedenfalls dann, wenn sie lediglich adaptierend verfährt und das Bewegtbild in allen seinen medialen Erscheinungsweisen in einer fraglosen Bild-Repräsentation aufgehen lässt, als sei die dargestellte Gegenständlichkeit unmittelbar als Zeichen zu verstehen, dem wie die Bedeutung dem Wort das alltagsweltliche Faktum als Sinn zugeschrieben werden kann.

Die Ablösung der Metapher von ihrer konkreten ästhetischen Ausformung in einer audiovisuellen Bildkomposition, deren Substituierung durch ein mehr oder weniger triftig ins Feld geführtes kognitives Schema, schließlich die Abweisung der wirkungsästhetischen Dimension rhetorischen Sprachgebrauchs zugunsten der kristallinen Klarheit einer von aller Emotionalität befreiten Verrechnung fixer Konstanten menschlichen Denkens – das alles verweist jedweden Aspekt medienästhetischer Modellierung und kulturhistorischer Situierung in den Bereich eines akzidentiellen Wissens, das der eigentlichen kognitiven Operation äußerlich bleibt. Das an der kognitiven Filmtheorie herauspräparierte Schema kehrt in der auf audiovisuelle Bildmedien angewandten Metaphernanalyse wieder: Je nachdem, was dem jeweils unterlegten kognitiven Schematismus (*image schema*, *mapping* oder *blending*) entgeht, es lässt sich immer als kulturelles Wissen der Rezipienten der rekonstruierten ‚eigentlichen‘ Bedeutung hinzuaddieren.

So gesehen beziehen sich die CMT-gestützten Analysen audiovisueller Metaphern häufig gar nicht auf die metaphorische Dimension audiovisueller Bilder – im vollen Sinne des audiovisuellen Bewegtbildes –, sondern auf ein vorab präpariertes Interpretandum, das die scheinbar selbstverständlich gegebenen, isolierten Aspekte der Bild-Repräsentation extrapoliert: ‚die Straße‘, ‚das Haus‘, ‚das Auto‘ werden für sich genommen, als handelte es sich um die entsprechenden Worte ‚Straße‘, ‚Haus‘, ‚Auto‘. Diese werden nicht auf ihre mediale Verfasstheit befragt, sondern unmittelbar als audiovisuelle Repräsentationen alltagsweltlicher Sachverhalte identifiziert. In dieser Herangehensweise werden audiovisuelle Bilder in der Folge weder kategorial von Texten oder Gemälden unterschieden, noch wird

zwischen unterschiedlichen medialen Formaten audiovisueller Bilder differenziert.

Der Rekurs auf universelle Schemata menschlichen Denkens gerät schließlich zur parodistischen Verkehrung exakter Wissenschaftlichkeit, wenn Untersuchungsgegenstände aus dem Feld der Kunst adressiert werden. Wird doch in der Regel auf eine neurowissenschaftliche oder kognitionspsychologische Forschung rekurriert, deren experimentell gesicherte Kenntnis kognitiver Schemata in ihren Abstraktionen weit davon entfernt ist, die konkreten Thesen zu stützen, die in solchen Analysen audiovisueller Bildmedien der angewandten *CMT* als Prämissen gesetzt werden.

Entsprechende Metaphernanalysen gleichen in ihren Reflexionen einem Museumsbesucher, der vor einem modernen Gemälde verharrt: Mit einem Schlag erkennt er in den Abstraktionen aus Farbflächen, Linien und Schraffuren ein Gesicht, das sich aus zwei, drei – oder sind es vier? – Gesichtern zusammensetzt. Berührt von dieser Entdeckung geht er weiter und sinniert darüber, welchen Sinn es haben mag, ein menschliches Gesicht derart deformiert und vielgesichtig aufgefächert zu malen. Allein gelassen mit seinem psychologischen Alltagswissen kommt er zu dem Schluss, es handle sich um eine Metapher, die eine multiple Persönlichkeit darstellen will; oder vielleicht handelt es sich doch eher um ein Bild der Identitätskrise des modernen Menschen: Das moderne Ich ist ein in Stücke gegangenes Porträt ...?

Nun könnte man sich über die Interpretation streiten. Tatsächlich ist sie nichts anderes als eben dies: eine Interpretation. Sie bezieht sich freilich auf Bild-Repräsentationen, die anderen Betrachtern erst gar nicht begegnen mögen. Sofern eine solche Interpretation aber mit der Algebra kognitiver Schemata begründet und beschrieben werden soll,<sup>24</sup> wird man noch einige Jahrhunderte neurowissenschaftlicher Forschung abwarten müssen, bis die Lücke zwischen empirisch belastbaren Beschreibungsformen und den spekulativen Medienanalysen angewandter Kognitionspsychologie geschlossen sein wird. Bis dahin bleiben Behauptungen vom Typ ‚Ich sehe ein Gesicht, gebaut aus vielen Gesichtern‘ interpretative Aneignungen, die sich durch den Rückgriff auf neurowissenschaftliche und kognitionspsychologische Empirie lediglich rhetorisch als wissenschaftlich gesicherte Erkenntnis verkleiden. Für die Begründung der Interpretation ist es immer noch naheliegender, den Gang durch die Gemäldegalerie fortzusetzen und die eigene Kunsterfahrung in der Geschichte der Poetik des Porträts zu situieren: Die nämlich kann uns tatsächlich viel von der Geschichte eines Ichs erzählen, das im Gang durch die Zeit

---

<sup>24</sup> Vgl. Todd Oakley: *From Attention to Meaning. Explorations in Semiotics, Linguistics, and Rhetoric*, Bern 2009.

sein Gesicht immer neu und anders entwarf. Womit nichts über die Möglichkeiten ausgesagt werden soll, die Prozesse subjektiver Aneignung im Kunstkonsum selbst zum Gegenstand wissenschaftlicher Analysen zu machen.

Der Betrachter befindet sich vor dem Gemälde wie vor dem Wittgensteinischen Kippbild: Er kann immer nur eines sehen, entweder den Hasen oder die Ente. Er sieht entweder den Hasen, die Bild-Repräsentation – ein vielgesichtiges Gesicht –, oder er sieht die Ente, das Bild selbst, die Farben, die Lineatur, die Komposition. Eine Analyse metaphorischer Übertragung in visuellen Darstellungen müsste nicht nur das eine im Verhältnis zum anderen in den Blick bekommen, um darin zu beschreiben, wie sich das betrachtende Auge einnistet in die Textur des Gemäldes und es zum Kippen bringt; sie müsste noch den Moment, in dem das Bild als Bild-Repräsentation greifbar wird, den Vorgang des Kippens selbst als das affektive Initial des metaphorischen Prozesses in die Analyse miteinbeziehen.

Denn gerade das momenthafte Ereignis, mit dem die Schraffuren, Farbflächen und Linien eines Bildes für einen Betrachter zu einem vielgesichtigen Gesicht zusammenschießen (ihn in seinem Wahrnehmen affizieren und in dem Affekt sein Denken in eine bestimmte, d. h. gerichtete Dynamik versetzen), bezeichnet das konstitutive Moment, das die Wahrnehmung der Erfahrung der eigenen Wahrnehmung vor dem Gemälde von der Alltagswahrnehmung unterscheidet. Was letztlich nichts anderes meint als eine konstitutive Differenz zwischen ästhetischer und alltagsweltlicher Erfahrung. Letztere kennt keine mediale oder ästhetische Zurückung auf eine bestimmte Dynamik des Wahrnehmungsprozesses hin.<sup>25</sup>

Übergeht man die Differenz zwischen der alltäglichen Wahrnehmung von Dingen und Handlungen und der von Gemälden, Fotos oder audiovisuellen Bewegtbildern, so erliegt man einem Missverständnis, das in Hinblick auf die Sprache undenkbar wäre: Würde man doch niemals dem Wort ‚Hase‘ die Bewegungsqualität zuschreiben, die den Hasen als Lebewesen zur Metapher für schnelles Laufen prädestiniert.

### **CMT und Metaphorologie**

Ginge man von der zitierten theoriegeschichtlichen Selbstverortung der CMT aus und bestimmte man diesen Ort tatsächlich durch den Bruch mit der Tradition von Rhetorik und Poetik, gäbe es keine Verbindung zwischen kulturwissenschaftlicher Medientheorie und kognitionswissenschaftlicher Metapherntheorie; ebenso

---

<sup>25</sup> Vgl. dazu Koch: Die Wiederkehr der Illusion, S. 44 sowie meine noch folgenden Ausführungen zu Dewey.

wenig gäbe es eine Vermittlungsmöglichkeit im Begriff der Metapher zwischen philosophischer Metaphorologie und kognitiver Metapherntheorie – die Vertreter der einen lägen aus Sicht der anderen schlicht und einfach falsch. Tatsächlich handelt es sich vorderhand um grundlegend verschiedene Gegenstandsbestimmungen dessen, was als Metapher überhaupt wissenschaftlich untersucht werden soll. Die eine betrifft einen psychologischen Sachverhalt, die andere die Kulturgeschichte des menschlichen Denkens. Was nichts anderes sein kann, als eine Geschichte der Entstehung und Veränderung der symbolischen Formen, in denen sich die Praxis menschlichen Denkens niederschlägt – seiner Aggregate, Medien, Artefakte und Sprachen, seiner Diskurse.

Folgte man konsequent der zitierten Position Forcevilles, wären alle Metaphern, die auf poetischen oder rhetorischen Konzepten fußen, nur soweit von Interesse, als sie auf einen letzten Grund in kognitiven Schemata verwiesen. Damit aber wären alle Aspekte metaphorischer Sinnkonstruktionen in Literatur und Kunst, die nicht in der bloßen Reproduktion kognitiver Schematismen aufgehen (was sie zumindest dem Common Sense aller Kunst- und Literaturtheorie folgend *per definitionem* nicht tun), aus dem Forschungs- und Anwendungsfeld der *CMT* ausgeschlossen. Eine solche Gegenstandsbestimmung mag Anspruch darauf erheben können, Erkenntnisse und Hypothesen der Kognitionspsychologie über kognitive Schematismen qua Exemplifizierung beispielhafter kultureller Produkte zu stützen. Sie kann keine Aussagen über die kulturelle Sinnproduktion selbst begründen.

Kurzum: Als deskriptive kognitionspsychologische Forschung mögen die genannten Ansätze ihren Sinn haben. In der Anwendung auf politische, kulturelle und medienästhetische Fragen führt die verengte Gegenstandsbestimmung zu höchst zweifelhaften Ergebnissen.

Der ausschließende Gegensatz zwischen *CMT* einerseits, Poetik und Rhetorik andererseits, lässt sich nur dann methodisch sinnvoll operationalisieren, wenn die Identifikation eines wie immer gearteten universellen kognitiven Schemas die Frage nach der metaphorischen Sinnkonstruktion – respektive deren politischer, ideologischer oder kultureller Funktion – bereits beantwortet würde. Eine recht waghalsige Vorstellung, die aber doch von einigen der Arbeiten zu audiovisuellen Metaphern suggeriert wird.<sup>26</sup>

---

**26** Forceville: *Metaphor in Pictures and Multimodal Representations*; Maarten Coëgnarts/Peter Kravanja: *Metaphor in Pictures and Multimodal Representations*, in: *The Cambridge Handbook of Metaphor and Thought*, hrsg. v. Raymond W. Gibbs, Cambridge 2008, S. 462–482; Coëgnarts/Kravanja: *From Thought to Modality*; Maarten Coëgnarts/Peter Kravanja: *On the Embodiment of Binary Oppositions in Cinema: The Containment Schema in John Ford's Westerns*, in: *Image [& Narrative* 15 (1), 2014, S. 30–42.



Das Problem also liegt in der *Anwendung* der *CMT*, ihrer Erkenntnisse und Hypothesen auf dezidiert kulturgeschichtliche oder gesellschaftspolitische Fragestellungen, die auf die Rekonstruktion kultureller Sinnressourcen, politischer Ideologien und affektiver Vergemeinschaftungsformen abzielen. Die Medienanalysen angewandter *CMT* gewinnen genau in dem Maße kulturwissenschaftliche Relevanz, wie sie die rhetorische und ästhetische Dimension audiovisueller Bilder einbeziehen. Die kognitiven Schemata werden dann nämlich zu einer Beschreibungssprache, welche die Dynamik des Prozesses metaphorischer Konzeptbildungen selbst in den Fokus rückt<sup>27</sup> – statt universale Konzepte im Rekurs auf physiologische Prozesse als ein und für allemal gegebene Strukturen zu identifizieren.

Vor dem Hintergrund der bisherigen Überlegungen kann man zunächst festhalten, dass die Modelle kognitiver Prozesse über die physiologischen oder psychologischen Grundlagen mentaler Aktivitäten belehren, nicht aber über die jeweils konkreten, historisch situierten kommunikativen Interaktionen, welche kulturelle Sinnbildungsprozesse prägen; seien diese nun als Face-to-Face-Kommunikation oder als medial vermittelte Operationen komplexer Gemeinschaften gedacht.

Demgegenüber ist in der im ersten Teil entwickelten phänomenologisch-medienästhetischen Perspektive die Geschichtlichkeit audiovisueller Bilder konstitutiv. Sie betrifft jeden Aspekt ihrer sinnstiftenden Funktion innerhalb medial strukturierter menschlicher Interaktion; die historische Dimension ist irreduzibel – und keineswegs durch den Rekurs auf allgemeine Grundlagen menschlichen Denkens zu umgehen. Stattdessen wäre in der Struktur des metaphorischen Prozesses der Zugang zu suchen, mit dem das Zusammenspiel historisch-kultureller Sinnhorizonte, medialer Repräsentationsmodi, kognitiver Schemata und einer unbestimmten *Pluralität* von Rezipienten in den Blick zu nehmen ist.<sup>28</sup>

### **Metaphern, mit denen wir leben**

Damit sind zwei auf den ersten Blick einander ausschließende Perspektiven aufgezeigt. Dennoch soll es im Folgenden mein Ziel sein, Bausteine einer Theorie der Metaphern audiovisueller Bilder zusammenzutragen, die sich zwischen diesen beiden Flanken zu bewegen sucht.

---

<sup>27</sup> So lässt sich etwa bei Kathrin Fahlenbrach beobachten, wie der *CMT* folgende Analysen audiovisueller Metaphern durchaus immer wieder auf Argumentationen zurückgreifen, die eindeutig der Rhetorik entliehen sind.

<sup>28</sup> Vgl. dazu das Unterkapitel *Kooperation und Kognition* dieser Studie (S. 124–S. 127).

Warum aber sollte man überhaupt versuchen, diese Sachverhalte aufeinander zu beziehen? Zunächst deshalb, weil die angewandte Metaphernforschung die audiovisuellen Medien als einen bevorzugten Gegenstand entdeckt hat. Dabei sind ihre Ergebnisse aber häufig unbefriedigend und fallen weit hinter die methodologischen Standards der Medienwissenschaft zurück. Die wechselseitige Unkenntnis wissenschaftlicher Fachexpertise muss man wohl als notwendige Folge der disziplinären Entropie des wissenschaftlichen Diskurses in Zeiten der Transdisziplinarität hinnehmen – was wäre denn auch die Alternative? Entscheidend ist deshalb, dass wir aus unserer eigenen Forschung zu Metaphern in audiovisuellen Medien<sup>29</sup> den Schluss ziehen, dass die gegensätzlichen Gegenstandsbestimmungen durchaus komplementäre Aspekte des Phänomenbereichs adressieren. Ist es doch ebenso widersinnig, mediale Ausdrucksformen menschlicher Interaktion durch den Rekurs auf ihre biologischen Grundlagen in ihrem Sinnpotential erfassen und interpretieren zu wollen, wie es umgekehrt nicht zu leugnen ist, dass jede medial strukturierte menschliche Interaktion Wesen einschließt, die einer biologischen Spezies angehören, über deren kognitive Fähigkeiten man sehr viele gut begründete Aussagen treffen kann.

Angesichts der geschilderten Ausgangslage mag es kurios anmuten, davon auszugehen, dass die *CMT* selbst einen gewichtigen Baustein zu einer Theorie kultureller Ausdifferenzierung des Metapherngebrauchs beisteuern könnte. Aber man sollte den Umstand, dass sich die *CMT* in ihren Anfängen auf den Nachweis stabiler semantischer Verknüpfungen im Hintergrund verschiedenster sprachlicher Ausdrücke stützte – statt die Hypothese solcher Verknüpfungen als Interpretation kultureller Artefakte ins Feld zu führen –, nicht wiederum als bloße Rhetorik abtun. Ein Indiz dafür mag die gegenwärtige Forschung zur Sprachentstehung liefern, die in der *CMT* einen ihrer initialen Impulse fand. Wenn diese Forschung heute die Sprachentstehung selbst als Differenzierung kognitiver Schemata denkbar werden lässt, die sich gerade keinem neuronalen Schematismus verdankt, sondern der menschlichen Fähigkeit zur Kooperation, d. h. spezifischer Formen menschlicher Interaktion, wird deutlich, wie sehr sich der ursprüngliche Impetus von Lakoffs und Johnsons *Metaphors We Live By* (1980), der Gründungsschrift der *CMT*, in der kulturwissenschaftlichen Anwendung ins Gegenteil verkehrt hat.

Tatsächlich nämlich nimmt die *CMT* ihren Ausgang in der Entdeckung der Metapher als eines Instruments des Denkens. Die Entdeckung ist selbst gleich-

---

<sup>29</sup> Müller/Kappelhoff: Cinematic Metaphor; Hermann Kappelhoff/Cornelia Müller: Embodied Meaning Construction. Multimodal Metaphor and Expressive Movement in Speech, Gesture, and Feature Film, in: *Metaphor and the Social World* 1 (2), 2011, S. 121–153.

sam der späte sprachtheoretische Abkömmling eines fundamentalen Bruches in der Geschichte der westlichen Philosophie. Mit ihr erreicht die Linguistik einen Begriff der Metapher, der seinen Ausgang in Nietzsches „Über Wahrheit und Lüge im außermoralischen Sinne“ (1873) hatte; ein Text, den man ohne weiteres als paradigmatisch werten kann für die fundamentale Umwertung, die das Verständnis von Wahrheit am Ausgang des 19. Jahrhunderts erfahren hat. Die Wahrheit wurde herabgestuft auf eine Beschreibung unserer gemeinschaftlich geteilten Wirklichkeit, die nun selbst wiederum keinen anderen Boden mehr fand als die rhetorische Kraft immer neuer Metaphernbildungen.<sup>30</sup> Die Metapher wurde mit Nietzsche zum Paradigma einer wirklichkeitszeugenden Rhetorik.<sup>31</sup> Damit aber war der Geltungsbereich von Poetik und Rhetorik auf alle Bereiche menschlicher Interaktion ausgeweitet.

Was im Ergebnis nichts anderes heißt, als dass die Gegenstandsbestimmung der *CMT* qua rigider Ausschlussrhetorik wohl zur strategischen Legitimation eines Forschungsansatzes, nicht aber zum Verständnis der Metapher im gegenwärtigen kulturwissenschaftlichen Diskurs beiträgt. Deshalb mag es hilfreich sein, sich noch einmal des historischen Ausgangspunktes der *CMT* zu vergewissern.

## 2.3 Kognitive Metaphertheorie und Cinematic Metaphor

Es ist Max Black, der 1977 in seiner Revision der Interaktionstheorie festhält, dass es ein entscheidender Schritt für die Metaphernforschung wäre, wenn sie metaphorische Aussageformen als Instrumente des Denkens konzipieren würde. Dann nämlich wäre die These nicht mehr obskur, die ihm einigen Widerspruch eingebracht hat – diese besagt, dass Metaphern neue Relationen in die Welt bringen können, die es vorher so nicht gegeben hat:

Wenn manche Metaphern das sind, was man ‚kognitive Instrumente‘ nennen könnte, die unerlässlich sind zur Wahrnehmung von Verbindungen, welche, einmal erkannt, *dann* wirklich vorhanden sind, dann wäre der Beispielfall für die These gefunden. Funktionieren Metaphern je als derartige ‚kognitive Instrumente?‘ Ich glaube ja.<sup>32</sup>

---

<sup>30</sup> Vgl. Sarah Kofman: Nietzsche und die Metapher, Schmalkalden 2014. In gewisser Weise ist das die Entdeckung Nietzsches für die Linguistik. Ich komme am Ende darauf zurück.

<sup>31</sup> Vgl. Andrew Hines: The Revolution of the Status of the Metaphor in Nietzsche’s Philosophy. Vortrag gehalten am 4. Juli 2016 auf der RaAM 11, FU Berlin.

<sup>32</sup> Max Black: Mehr über die Metapher, in: Theorie der Metapher, hrsg. v. Anselm Haverkamp, Darmstadt 1996, S. 379–413, hier S. 408–409.

## Die Metapher als Paradigmenwechsel der Sprachtheorie

Interessant ist die These auch und vor allem mit Blick auf den damit verbundenen Umkehrschluss: Läuft dieser doch letztlich darauf hinaus, dass die gemeinsame Wirklichkeit genauso weit verfügbar ist, wie es Metaphern gibt, die diese Wirklichkeit beschreiben können.

Versucht man Blacks Argumentation rückblickend zu bewerten, wird deutlich, dass seine Argumente in ihrer inneren Logik bereits auf den Bruch mit dem dominanten sprachtheoretischen Paradigma zulaufen. Seine Überlegungen zur Dynamik der metaphorischen Interaktion – Hartmut Winkler hat dies in einem Aufsatz von 1998 präzise herausgearbeitet<sup>33</sup> – sprengen den Rahmen eines strukturalistisch definierten Sprachsystems, das dem Denken notwendig vorhergeht. Denn Metaphern weisen auf sprachliche Prozesse zurück, die in der Spontanität menschlicher Akteure gründen. Sie implizieren gleichermaßen subjektive Kreativität und historische Sedimentierungsprozesse – sie implizieren Prozesse der Sprachentstehung. Den Riss, der sich damit inmitten eines von Strukturalismus und Semiotik geprägten sprachtheoretischen Diskurses öffnet, gilt es im Auge zu behalten.

Blacks Überlegungen verweisen mithin auf drei grundlegende Aspekte, die zentral sind für die Beziehung zwischen audiovisuellen Bildern und Metaphern:

1. Metaphern sind Werkzeuge des Denkens;
2. sie beziehen sich gleichermaßen auf die sinnlich wahrnehmbare Wirklichkeit und deren symbolische Struktur;
3. diese Wirklichkeit ist nicht objektiv gegeben; sie ist vielmehr stets bezogen auf bestimmte Perspektiven und Aspekte ihrer Beschreibung.

Alle drei Aspekte markieren letztlich die Bruchstelle der bis dahin vorherrschenden Sprachtheorien. Doch erst der Schluss, den Lakoff und Johnson aus den Einsichten Blacks u. a. zogen, vollendet den grundlegenden Paradigmenwechsel:

The general principles [davon, wie Menschen ihre Erfahrungen verstehen, H.K.] involve whole systems of concepts rather than individual words or individual concepts. We have found that such principles are often metaphoric in nature and involve understanding one kind of experience in terms of another kind of experience.<sup>34</sup>

<sup>33</sup> Winkler: Metapher, Kontext, Diskurs, System.

<sup>34</sup> Lakoff/Johnson: Metaphors We Live By, S. 116–117.

Nun sind es nicht mehr die Strukturen oder generativen Faktoren von Sprachsystemen, die das menschliche Denken präfigurieren, sondern Systeme von kognitiven Schemata, die das Denken als Verknüpfungs- und Differenzierungsoperation beschreiben, aus der die Sprachbildungen selbst noch hervorgehen. In der Gegenwart der Metaphern im allgemeinen Sprachgebrauch ist ein Prinzip solcher kognitiver Operationen greifbar, dem ein umfassender paradigmatischer Wert zugesprochen wird: die bedeutungsgenerierende Interaktion zwischen distinkten Erfahrungsbereichen, „understanding one kind of experience in terms of another kind of experience.“

Dem sprachtheoretischen Paradigmenwechsel folgend, lassen sich sehr unterschiedliche Forschungsperspektiven ableiten. Ich erwähne jene drei, die heute das Feld der kognitiven Metaphernforschung bestimmen:

1. Man kann die Metaphernkonzepte verschiedener Kulturen, Medienformate und Kommunikationsformen untersuchen, um theoretische Modelle basaler (d. h. physiologisch fundierter) Prozesse des menschlichen Denkens zu erschließen – und diese Modelle umgekehrt zur Grundlage empirischer Studien über die Verfasstheit des menschlichen Denkens machen. Man kann das als die kognitionspsychologische Perspektive bezeichnen.
2. Andererseits kann man diese Art theoretischer Modelle nutzen, um kognitive Prozesse oder metaphorische Operationen in sprachlichen Äußerungen, Texten oder Medienformaten verschiedenster Art zu identifizieren und zu erklären, indem man diese Konzepte auf allgemeine Prinzipien (z. B. Bildschemata) zurückführt. Auf diese Weise lassen sich kulturelle Phänomene unter einem kausalen Erklärungsmodell subsumieren; dem schier endlosen Spiel der permanenten Modulation dessen, was wir als gemeinschaftlich geteilte Wirklichkeit verstehen, kann so in den Prinzipien menschlicher Kognition ein Boden universell gültiger Tatsachen eingezogen werden. Die im vorherigen Kapitel beschriebene Applizierung der *CMT* auf beliebige kulturelle Gegenstände verfährt regelmäßig in dieser Weise.
3. Schließlich können die theoretischen Einsichten zur metaphorischen Operation und deren Konzepte als Ausgangspunkt dienen, um die Dynamiken medialer Interaktion und kulturhistorischer Prozesse – und deren Situiertheit in konkreten sozialen Interaktionen – zu untersuchen. Und zwar unter der Maßgabe, dass solche Operationen im Grenzfall durchaus unsere Vorstellung von Wirklichkeit verändern, indem sie neue Beschreibungen oder spezifische Perspektivierungen der Wirklichkeit hervorbringen.

Ein nicht unerheblicher Teil der Probleme, die in der Anwendung der *CMT* auf audiovisuelle Bilder entstehen, rührt von der mangelnden Differenzierung solcher Geltungsansprüche her. So sollte an der Kritik der angewandten Meta-

phernforschung deutlich geworden sein, dass Fragen, die sich auf die Analyse kultureller Sinnpotentiale audiovisueller Bilder beziehen, vor allem der dritten und letzten Variante zuzuordnen sind.

Wie aber lässt sich das kognitionstheoretische Verständnis von Metaphern für eine Analyse audiovisueller Bilder fruchtbar machen, die auf die Dynamiken medialer Interaktionen in ihren wirklichkeitskonstitutiven Aspekten abzielt? Dieser Frage möchte ich mich zunächst über eine kleine analytische Skizze annähern.

### **Das Klatschen der Hände: REAR WINDOW**

In der Sequenz, um die es im Folgenden gehen soll, können wir beobachten, wie eine der berühmtesten Nebendarstellerinnen des klassischen Hollywoodkinos – Thelma Ritter als Krankenschwester Stella in Alfred Hitchcocks REAR WINDOW (DAS FENSTER ZUM HOF, USA 1954) – einem Patienten – dem Protagonisten des Films, Jeff (James Stewart) – ihre Vorstellung von Ehe und ihre Einstellung zur Liebe nahezubringen sucht. Sie tut dies vor dem Hintergrund, dass Jeff offenbar schon seit geraumer Zeit zaudernd und zweifelnd die Frage erwägt, ob er seine Freundin Lisa (Grace Kelly) nun heiraten soll oder nicht. Stella greift in dem Gespräch, bei dem sie den Patienten nicht nur mit ihren Worten, sondern auch mit ihren Händen massiert, und zwar den Rücken, zu einer Metapher:

REAR WINDOW (00:10:40 bis 00:11:10)

STELLA: Look, Mr. Jefferies. I'm not an educated woman. But I can tell you one thing – when a man and a woman see each other, and like each other – they oughta come together – wham, like a couple of taxis on Broadway. Not sit around analyzing each other like two specimen in a bottle.

JEFF: There's an intelligent way to approach marriage.

STELLA (scoffing): Intelligence! Nothing has caused the human race so much trouble as intelligence. Modern marriage!

In der Szene lässt weder die Bedeutung der Worte noch eine zeichenhafte Geste die Metapher entstehen. Vielmehr ist es die Art und Weise, wie eine Geste im Rhythmus der Inszenierung, des Schauspiels und des Sprechens eingebettet und arrangiert wird, die über den Verlauf der gesamten Szene hinweg buchstäblich Sinn macht: *Sich-Verlieben ist ein Unfall, ein Zusammenprall zweier Autos!* Mühelos lässt sich verstehen, worauf Stella hinaus will. Doch wenn wir ihre ver-

balen Äußerungen betrachten – „when a man and a woman see each other ...“ –, stellen wir fest, dass weder von ‚Verlieben‘ noch von einem ‚Autounfall‘ die Rede ist. Warum also drängt sich uns diese Metapher auf, wenn wir Stella zuhören und zuschauen?

Die Antwort auf diese Frage scheint recht einfach zu sein: Wir verstehen die Metapher, weil wir in der Lage sind, zwei verschiedene Erfahrungsbereiche miteinander zu verbinden: den Straßenverkehr in einer Großstadt und Liebesbeziehungen. Um auf diese Verbindung zu kommen, müssen wir, die Zuschauer, zunächst die verschiedenen Dimensionen des audiovisuellen Bildes zusammenbringen: den Schauspielstil, die wörtliche Bedeutung des Dialogs und ein verkörpertes Gefühl für den Rhythmus der Komposition des Bewegungsbildes.

Betrachten wir zunächst das Ineinander von Dialog und Geste. Sofort fällt auf, dass der verbale Ausdruck „wham, like a couple of taxis on Broadway“ von einer starken Geste vervollständigt wird. Der Ausruf „wham“ unterstreicht das Ineinanderklatschen der Hände Stellas. Für sich genommen verweist auch das Klatschen nicht unmittelbar auf einen Autounfall. Wohl aber fungiert es im Zusammenhang der *Mise en Scène* als eine Metonymie. Und dies in zweierlei Hinsicht.

Zum einen ist da der Klang des Klatschens, der als rein akustische Referenz das Krachen des Autounfalls evoziert. Zum anderen ruft die Bewegung des beidhändigen Klatschens den Eindruck auf, dass zwei antagonistische Kräfte aufeinanderprallen. Beide metonymischen Relationen werden also vor allem über unser sinnliches Erleben miteinander verknüpft und weniger über die ikonische Repräsentation oder die Wortbedeutungen. Ein verschwindend kleiner metonymischer Bezug (das Klatschen zweier Hände) drängt zur Imagination der großen dramatischen Szene: Autounfall.

Man könnte ein solches Erleben mit einem Begriff von Ray Gibbs als „embodied simulation“<sup>35</sup> bezeichnen: Der Zuschauer entwickelt ein Gefühl, ein körperliches Empfinden für das Klatschen der Hände. Wird das ‚Gefühl für‘ mit dem Sinn der Worte „two taxis on Broadway“ verbunden, ist der Quellbereich (*source domain*) einer metaphorischen Übertragung etabliert, die letztlich in der ästhetischen Erfahrung der filmischen Inszenierung gründet: die Kollision zweier gegenläufiger Kräfte, der frontale Zusammenprall zweier Autos.

Selbstredend mag das semantische Feld des sprachlichen Ausdrucks „a couple of taxis on Broadway“ noch eine ganze Reihe anderer metaphorischer

---

35 Raymond W. Gibbs: *Metaphor Interpretation as Embodied Simulation*, in: *Mind & Language* 21 (3), 2006, S. 434–458; vgl. S. 436: „Simply put, one reason why people interpret many verbal metaphors through embodied simulations is because this metaphoric language is rooted in bodily processes that people may imaginatively recreate during their ordinary use of such language.“

Übertragungen bereithalten. Ich möchte mich hier allerdings auf jene beschränken, die sich durch eine deskriptive Analyse der Inszenierung des filmischen Bildraums einholen lassen.

Doch was ist mit dem Zielbereich der Metapher? Der Dialog beinhaltet die Formulierung „a man and a woman seeing each other“ und „marriage“, aber das Konzept des Sich-Verliebens wird nicht ausgesprochen. Warum also scheint es so eindeutig zu sein, dass die Metapher, die hier gestaltet wird, ‚Sich-Verlieben ist wie ein Autounfall‘ lautet? Woher bekommen wir als Zuschauer die Idee, dass der Autounfall einen ganz bestimmten Aspekt des Erfahrungsbereichs Liebe aufruft: die Plötzlichkeit, das Unvorhersehbare, das Überwältigende des Verliebense? Die Antwort fällt ähnlich aus wie jene bezüglich des Quellbereichs: Nur dass es sich diesmal um den Gestus der Komposition des audiovisuellen Bildes der filmischen Inszenierung handelt. Es ist die Geste des audiovisuellen Bildes selbst, die hier zur körperlichen Simulation, zu einem ‚Gefühl für‘ wird.

Man kann zunächst einmal das Schauspiel von Thelma Ritter hervorheben. Die Art und Weise, wie sie sich durch den Raum bewegt, der schnelle, präzise Rhythmus ihres Sprechens – all das evoziert die Erfahrung kraftvoller Vitalität. Diese Erfahrung ist zwar durch das Schauspiel strukturiert, doch gewinnt sie zu einem erheblichen Teil erst durch die Rahmung der filmischen Inszenierung ihre Prägnanz. Das schnelle und pointierte Hin und Her des Dialogs wird durch den Rhythmus der Schnitte akzentuiert, und die zielgenauen Bewegungen der Schauspielerin durch den Raum werden mit den Schnitten und horizontalen Kamerabewegungen synchronisiert (siehe Abb. 11, Farbbogen 2).

Auf diese Weise entsteht für uns Zuschauer die Erfahrung, unmittelbar im eigenen Sehen und Hören an der Vitalität der Figur teilzuhaben. Die klatschenden Hände vermitteln die sinnliche Erfahrung der Kollision gegenläufiger Kräfte; aber im Rhythmus der Inszenierung durchdringt die Erfahrung kraftvoller Vitalität (als verkörpernde Simulation der Bildkomposition selbst) die gesamte Sequenz. Die Rezeption des kinematografischen Bildes erzeugt ein Gefühl, das seinerseits die Grundlage des metaphorischen Konzepts ‚Sich-Verlieben ist wie ein Autounfall‘ abgibt.

Die verkörpernde Simulation der Rhythmik der filmischen Komposition fügt dem „a man and a woman“ und „marriage“ ein ‚Gefühl für ...‘ hinzu: die Erfahrung, sich in einem kraftvollen und vitalen Rhythmus zu bewegen. Die Übertragung zwischen den zwei explizit genannten Erfahrungsbereichen „a couple of taxis on Broadway“ und „a man and a woman“ vollzieht sich tatsächlich als Interaktion; freilich als Interaktion von zwei verkörperten Simulationen, einem Gefühl für die Kollision und einem für die kraftvolle Vitalität. Erst in der wechselseitigen Durchdringung beider Erfahrungsbereiche (Autounfall, Liebesbeziehung) entsteht die Vorstellung von der Plötzlichkeit, Unvorhersehbarkeit und



dem Überwältigt-Werden der Verliebtheit; eine Metapher, deren innere Struktur auf der ästhetischen Erfahrung der filmischen Inszenierung, ihrer verkörpernden Wahrnehmung gründet.

Die Szene zwischen Stella und Jeff entfaltet also in zweifacher Hinsicht ein metaphorisches Konzept: Zum einen wird eine audiovisuelle Geste gestaltet, die die Vorstellung eines Autounfalls als Quellbereich etabliert: Ein Autounfall, das heißt, zwei antagonistische Kräfte prallen mit lautem Knall gegeneinander. Zum anderen wird diese spezifische verkörperte Simulation – die eines Autounfalls – in eine multidimensionale Zeitgestalt des audiovisuellen Bildes eingebettet.

Hitchcock nutzt alle Mittel audiovisueller Inszenierung, um zu einer filmischen Metapher zu kommen, die weder in der Rede des Dialogs geäußert wird, noch an einer einzelnen Geste festzumachen ist. Stattdessen bezieht er sich in der Inszenierung der Szene auf die physisch-sinnliche Response affektiv involvierter Zuschauer und auf die temporale Gestalt des kinematografischen Bildes – sodass im zeitlichen Verlauf der sich entwickelnden Szene der semantische Komplex der Phrasen „a couple of taxis“ und „a man and a woman“ hin zur Metapher ‚Sich-Verlieben ist wie ein Unfall, ein Zusammenprall zweier Autos‘ umgeformt wird. Die Metapher, die an keiner Stelle ausgesprochen wird, stellt sich nur im Wahrnehmen der Zuschauer her, sofern in ihrem ästhetischen Erleben der Sequenz zwei konkrete Erlebensqualitäten (Vitalität und Kollision) aufgerufen und miteinander verwoben werden.

Doch geht es mir in dieser analytischen Skizze nicht in erster Linie um die Sinnkonstruktion des Films als solchem, noch um die Rekonstruktion einer Metapher; vielmehr wollte ich die Interaktion herausarbeiten, die sich zwischen unterschiedlichen Erfahrungsbereichen, aber auch zwischen audiovisuellem Bild und Zuschauer vollzieht; ich wollte zeigen, dass beide Ebenen der Interaktion, unauflösbar ineinander verschränkt, integrale Aspekte der metaphorischen Sinnkonstruktion ausmachen.

Erst in der Interaktion zwischen Zuschauer und audiovisuellem Bild, erst in der Poiesis des Filme-Sehens entsteht aus dem Gewebe metaphorischer Übertragungen eine fiktionale Welt, als welche das Gewebe in allen Einzelheiten schließlich Sinn macht. Man nimmt den Faden metaphorischer Übertragungen dort auf, wo er einem sozusagen in die Hand fällt; selbst wenn damit nicht einmal die zentrale, die diskursstrukturierende Metapher<sup>36</sup> (in REAR WINDOW eben das Hinterhof-

---

<sup>36</sup> Vgl. Lynne Cameron zum Konzept der diskursstrukturierenden Metapher: Lynne Cameron: The Discourse Dynamics Framework for Metaphor, in: *Metaphor Analysis. Research Practice in Applied Linguistics, Social Sciences and the Humanities*, hrsg. v. Lynne Cameron, Robert Maslen, London 2010, S. 77–96; Lynne Cameron/Robert Maslen/Zazie Todd/John Maule/Peter Stratton/

fenster) erfasst ist, führt er dann doch hinein in das Gewebe eines Netzes immer weiterer Metaphorisierungen.

### Engagierte Zuschauer

Man mag einwenden, dass die kleine Sequenzanalyse zu *REAR WINDOW* in Gänze nichts anderes sei als eine eigensinnige Interpretation; stützt sie sich doch lediglich auf gewisse Evidenzen dessen, was als ästhetisches Erleben des Bewegungsbildes beschreibbar und deshalb mitteilbar ist. Auch die herauspräparierte Metapher kann einer Definition metaphorischer Konzepte vom Typ ‚a ist b‘ nicht genügen – denn der von mir zitierte Ausschnitt ist willkürlich gewählt aus einem Prozess gleichgerichteter metaphorischer Übertragungen, der sich über das Ganze des Films erstreckt. Ich würde solchen Einwänden nicht widersprechen. Betreffen sie doch allesamt Bestimmungen eines Typus metaphorischer Konzepte, der sich vom gängigen Schema angewandter kognitiver Metapherntheorie im selben Maße unterscheidet, wie er für das Verständnis audiovisueller Metaphern grundlegend ist.

Lässt man den ästhetisch involvierten und reflektierenden Zuschauer außer Acht – d. i. die Wahrnehmung von Zuschauern, die sich in die Interaktion mit dem audiovisuellen Bild hineinbegeben haben und die in dieser Interaktion mit dem Bewegungsbild den Horizont eines spezifischen subjektiven Erlebens des Bildes produzieren –, wird man nur einen Mann mit einem gebrochenen Bein sehen, der in einem Zimmer, gefesselt an seinen Rollstuhl, vor dem Hinterhofzimmer hockt und vor allem eines wünscht: Es möge da im Hof etwas geschehen, das die Langleweiligkeit erträglich macht! Aus dem Dialog mit der Krankenpflegerin wird man erfahren, dass er eigentlich einen aufregenden Beruf (erfolgreicher Fotoreporter) hat und eine schöne Geliebte, die er nicht unbedingt heiraten will. Man wird seinem gelinde perversen Vergnügen folgen und die Nachbarn im Hinterhof ausspionieren. Man wird amüsiert zuschauen, wie die beiden Frauen hineingezogen werden in die Faszination des Voyeurs; und dann wird man, wie sie, dem rätselhaften Verhalten des Mannes von Gegenüber nachsinnen, der ganz offensichtlich einen Mord begeht, um seine Ehefrau loszuwerden (siehe Abb. 12, Farbbogen 2).

---

Neil Stanley: The Discourse Dynamics Approach to Metaphor and Metaphor-Led Discourse Analysis, in: *Metaphor and Symbol* 24 (2), 2009, S. 63–89. Eine solche Metapher ist im Titel ‚Rear Window‘ vorgegeben, mit der der Erfahrungshorizont des Protagonisten auf die der Kinoszuschauer bezogen wird. Vgl. Anne Eusterschulte: *Blumenberg's Metaphorology and Hitchcock's Rear Window*, in: *Cinematic Metaphor in Perspective. Reflections on a Transdisciplinary Framework*, hrsg. v. Sarah Greifenstein, Dorothea Horst, Thomas Scherer, Christina Schmitt, Hermann Kappelhoff, Cornelia Müller, Berlin/Boston 2018.

Für involvierte Zuschauer freilich, die sich in ihren Empfindungen und Gedanken eingenistet haben in den Bildraum von REAR WINDOW und seiner Entfaltung bis zum Finale folgen, wird sich erst von diesem Ende her ein sicheres Gefühl für das Ganze der filmischen Welt einstellen. Und erst aus dem Gefühl für das Ganze der Situation, die sich in der leibhaften Gegenwart ihrer Wahrnehmungsempfindungen genau in dem Maße auffaltet, wie sie sich affizieren und in das sich entfaltende Bewegungsbild verwickeln lassen – erst aus diesem Gefühl für das Ganze der Situation und für das eigene Verwickelt-Sein erschließt sich die Metapher als ein komplexer interaktiver Prozess zwischen Bewegungsbild und rezipierenden Körpern.

Solchermaßen in die dynamisch sich entwickelnde Situation verwickelte Zuschauer werden die oben skizzierte Metapher tatsächlich erst in den vielfältigen metaphorischen und metonymischen Verästelungen und Überlagerungen greifen, die endlich in einem weiteren Unfall des Protagonisten münden: Er stürzt aus eben jenem Hinterhoffenster, das involvierte Zuschauer sehr bald schon als Metapher ihres eigenen Wahrnehmungserlebens verstehen mögen. Sind sie doch ganz ähnlich wie der Protagonist im Rollstuhl in ihrer Bewegungsfreiheit massiv eingeschränkt; haften ähnlich wie dieser allein mit ihren Augen am Screen, mehr oder weniger begierig auf der Suche nach etwas *Sex and Crime*, oder vielleicht doch eher auf eine romantische Liebesgeschichte oder melodramatische Tränenseligkeit hoffend. Taktisch engagierte Zuschauer werden dies alles bekommen: Erotik, Thrill, Suspense, Horror, Melodramatik und romantische Liebe; die Modi des Genrekinos werden im Verlauf des Films so gegensätzlich und konfliktreich aufeinander bezogen wie das Verlangen nach Bewegungsfreiheit und die verrippten Beine des Protagonisten. Sie werden in den dramatischen Kollisionen der affektiven Modi des Hollywood-Genrekinos durch ein Wechselbad von Gefühlen geführt, das ihnen als Gefühl für die filmische Welt den Zugang zu den Ambivalenzen, Widerständen und Abwehrpositionen eröffnet, die von Anfang an als innerer Konflikt der *dramatis personae* angezeigt werden: Der Wunsch nach individueller Selbstbestimmung und die Sehnsucht nach Gemeinschaft – gleichviel, ob man diese nun Ehe oder Liebe nennt.

Ein dergestalt in den Verlauf des Films verwickelter Zuschauer mag erst am Ende, erst vom Fenstersturz des Helden her betrachtet die Metapher des Auto-unfalls mit der Idee des Sich-Verliebens komplementieren: Nämlich dann, wenn der Protagonist, als er sich mit zwei gebrochenen Beinen jeder Fähigkeit zur Flucht oder auch nur zur autonomen Bewegung beraubt sieht, endlich in die Ehe eingewilligt zu haben scheint, erschöpft und müde sich ganz der Obhut der Hollywood-Schönheit überantwortet hat. Einem solchen Zuschauer mag von der finalen Apotheose her eine letzte Verstrebung in der metaphorischen Konstruktion zuwachsen, welche die resolute Pflegerin in der Rhetorik der amerikani-

schen Screwball Comedy auf den Weg eines langen Prozesses der Elaborierung des metaphorischen Konzepts über immer neue Verzweigungen, Kreuzungen und Vertiefungen mit anderen metaphorischen Bezügen brachte, während sie dem so ungeduldigen wie hilflosen Patienten den nackten Rücken massierte.

Diese letzte Verstrebung meint den Prozess der affektiven Verwicklung der Zuschauer in Gänze, ihres zunehmenden affektiven Engagements, das schließlich in ein Gefühl für das Selbstgefühl des Protagonisten mündet und dergestalt erst dem Gefühl für das Ganze der fiktionalen Welt des Films retrospektiv eine gewisse Stabilität verleiht. Auf die fiktionale Wirklichkeit von REAR WINDOW bezogen ist noch die Spannung zwischen sozialer Bindung (Ehe) und individuellem Freiheitsstreben eine Metapher, die einen Topos des amerikanischen Genrekinos aufruft und als generische Form des Gendering greifbar werden lässt: die Frage nach den individuellen Möglichkeiten und sozialen Grenzen eines selbstbestimmten Lebens.

### Die Cinematic Metaphor

Liebe ist ein höchst abstraktes Konzept, und es ist nicht wirklich klar, was mit den Worten Verliebtsein und Sich-Verlieben verbunden ist. Was wir unter Liebe oder Verliebtsein letztlich verstehen, kann schwerlich außerhalb eines gegebenen sozialen oder kulturellen Kontextes entschieden werden. Und selbst dann ist der Sinn des Ausdrucks abhängig von den sozialen Gepflogenheiten spezifischer kultureller Subgemeinschaften und konkreten individuellen Erfahrungen. Insofern mag das metaphorische Konzept, das ich herauszustellen suchte, durchaus die Funktion haben, so etwas wie ein gemeinschaftliches Verständnis der Zuschauer für eine höchst idiosynkratische, von Paranoia und Misstrauen geprägte Sichtweise auf die Liebe herzustellen; eine Sichtweise, die sich den Zuschauern als Erfahrung von REAR WINDOW vermittelt, indem die filmische Inszenierung den Sinn an die Sinnesqualitäten konkreter filmischer Bildräume bindet.

Lakoff und Johnson mögen solche Komplikationen mit dem gar nicht so selbstverständlichen Verstehen allgemeinverständlicher Worte, die wir gebrauchen, um uns auf eine gemeinsam geteilte Wirklichkeit zu beziehen, im Blick gehabt haben, als sie folgende Beobachtung beschrieben: „The metaphors come out of our clearly delineated and concrete experiences and allow us to construct highly abstract and elaborate concepts [...]“<sup>37</sup>

---

37 Lakoff/Johnson: *Metaphors We Live By*, S. 105.

Ihre Überlegung impliziert einige Vorannahmen, die für die Analyse audiovisueller Bilder von großer Bedeutung sind – auch wenn sie in der angewandten Metaphernforschung häufig in den Hintergrund treten:

1. Metaphorische Konzepte sind keine fixen bedeutungsgenerierenden Schemata, sondern *operative Schemata, die aus Denk- und Verstehensproblemen erwachsen*.
2. Metaphorische Konzepte beziehen sich immer *sowohl auf die ästhetischen Aspekte* (Wahrnehmungskonstellationen), d. h. die sinnlich-perzeptive Dimension bestimmter Erfahrungsdomänen („clearly delineated and concrete“), als auch *auf abstrakte semiotische, begriffliche oder textuelle Konstrukte* („highly abstract“).
3. Die metaphorischen Konzepte haben *die pragmatische Funktion, verstandesmäßig oder sozial unzugängliche Erfahrungsdomänen durch Analogiebildungen zu erschließen*. Sie beziehen sich auf die Grenzen des allgemein geteilten Wissens, auf das Unbekannte und Unverstandene.

Mit Blick auf die analytische Skizze zu REAR WINDOW lässt sich nun resümierend ein Typus metaphorischer Konzepte bestimmen, der für das Verständnis audiovisueller Bewegtbilder paradigmatisch ist. Wir sprechen mit Blick auf diesen Typus von *Bewegungsbild-Metapher* oder – um einen englischsprachigen Terminus zu verwenden – von der *Cinematic Metaphor*.<sup>38</sup> Der Begriff umfasst, wie ich in der filmanalytischen Skizze zu zeigen suchte, alle Wechselbeziehungen zwischen dynamisch-energetischen (den Typus von Kräften und deren Wirkung betreffend) und kinematischen Aspekten (Bewegungsfigurationen in gegebenen zeitlich/räumlichen Relationen). D. h., er bezieht sich vor allem auch auf die dynamischen Veränderungen der Zeit/Raum-Verhältnisse als energetische Muster. Analog zum Terminus ‚filmisches Bewegungsbild‘ beziehen wir uns mit dem Begriff einerseits auf das audiovisuelle Bild als ‚multidimensionale Gestalt eines Erfahrungskomplexes‘,<sup>39</sup> die erst im Durchgang der Rezeption als ein Bild der Bewegung (ein Bild der dynamischen Veränderbarkeit von Bewegungsfigurationen) greifbar wird; andererseits ist damit die Metapher selbst als Produkt einer Poiesis des Filme-Sehens verstanden, als das Ergebnis eines Herstellungsprozesses, der die Rezeption des audiovisuellen Bildes selber meint.

---

<sup>38</sup> Wir wollen hier den englischen Begriff beibehalten, da er sich in seiner Spezifität kaum sinnvoll (oder nur unzureichend mit ‚filmische Metapher‘) ins Deutsche übersetzen lässt. Darüber hinaus wollten wir den technischen Terminus ‚kinematisch‘ nicht verwenden.

<sup>39</sup> Lakoff/Johnson, S. 81.

Die Rekonstruktion einer Bewegungsbild-Metapher zielt daher auch nicht auf die Identifikation einer einzelnen Bedeutungsrelation. Sie ist vielmehr selbst als Bildraum zu begreifen: D. h., sie meint das Netz metaphorischer Verweise zwischen den Dingen, Akteuren und Aktionen, deren innere Gesetzmäßigkeit sich allein im Spiel dynamisch ineinandergreifender Übertragungen unterschiedlichster Erfahrungsbereiche als poetische Logik der Verknüpfung bekundet. Die Sinnkonstruktion einer *Cinematic Metaphor* ist also immer ein dynamischer Bildraum in seinen affektiven, perzeptiven und konzeptuellen Dimensionen – wie das Zimmer mit dem ‚Fenster zum Hof‘, das den Blick auf die vielen anderen Fenster gegenüber freigibt, hinter denen sich Szenen des alltäglichen Lebens präsentieren, die in der Verkörperung des Voyeurs (der Protagonist, die Zuschauer) zu einem dynamischen metaphorischen Geflecht sich verbinden.

Ich möchte die Überlegungen zur *Cinematic Metaphor* knapp rekapitulieren:

1. *Cinematic Metaphors* bezeichnen solche Metaphern, die den Fiktionalisierungsprozess von Rezipienten audiovisueller Bilder strukturieren; gleichviel, ob es sich um Nachrichtenbeiträge, Dokumentationen oder Spielfilme handelt.
2. Tatsächlich legen alle unsere filmanalytischen Studien nahe, dass metaphorische Konzepte in audiovisuellen Bildern nicht als statische Einheiten zu fassen sind, die sich gemäß dem Schema ‚a ist b‘ formulieren lassen. Vielmehr sind sie als Bildräume zu rekonstruieren, in denen unterschiedliche Erfahrungsbereiche in einem dynamischen, d. h. in der Zeit sich entfaltenden Geflecht ineinandergreifender metonymischer und metaphorischer Übertragungen und Interaktionen miteinander in Beziehung treten<sup>40</sup> – ein Netz von Bezügen, das sich als dynamischer Prozess über das Ganze eines Films, Videos etc. ausbreitet.<sup>41</sup>
3. In der Konsequenz bedeutet dies, dass die Analyse des Sinns audiovisueller Bilder immer nur vermittelt über die Rekonstruktion des Prozesses der Entfaltung des filmischen Bewegungsbildes im sinnlichen Erleben konkreter Rezipienten, d. i. die ‚multidimensionale Erfahrungsgestalt‘ des Ganzen eines Films, Videos, Werbeclips oder Nachrichtenbeiträge, zu leisten ist. (Im ersten

---

<sup>40</sup> Vgl. Hermann Kappelhoff/Sarah Greifenstein: Metaphorische Interaktion und empathische Verkörperung: Thesen zum filmischen Erfahrungsmodus, in: Empathie im Film. Perspektiven der Ästhetischen Theorie, Phänomenologie und Analytischen Philosophie, hrsg. v. Malte Hagener, Ingrid Vendrell Ferran, Bielefeld 2017, S. 167–193.

<sup>41</sup> Nur, wenn wir einzelne audiovisuelle Repräsentationen aus ihrem kompositorischen Zusammenhang lösen, lassen sie sich als isolierte Metaphern vom ‚a ist b‘-Typus identifizieren. Aber dann verhalten sie sich bezogen auf das metaphorische Gefüge des Ganzen des filmischen Bildes indifferent.

Teil des Buches habe ich eine solche Analyse mit dem Begriff des Bildraums verbunden.) Erst auf dem Boden der Rekonstruktion des Prozesses ästhetischen Erlebens kann die Analyse sich auf vorab bestehende semiotische Systeme, kulturelle Semantiken oder kognitive Verstehensmodelle beziehen. Die Rahmung, innerhalb derer sich der Prozess der *Sinnbildung* entfaltet, muss im je besonderen Fall zunächst als mediale und ästhetische Rahmung rekonstruiert werden, bevor der durch sie präformierte Verstehensprozess überhaupt analysierbar ist.

4. Zumindest für die künstlerischen Formen audiovisueller Bilder gilt, dass eine Analyse sich immer mit dem Problem der Mittelbarkeit des ästhetischen Erlebens auseinandersetzen muss. Denn letztlich ist der einzige Zugang zu diesem Erleben eine Kenntnis der subjektiven Taktik der Aneignung. Man ist überwältigt, gerührt, bewegt von einer Performance, Videoinstallation, einem Film etc., aber es fällt schwer, objektiv verhandelbare Gründe anzugeben. Für ästhetische Urteile gilt im Besonderen, was im Letzten auch auf die Metapher zutreffen mag: Sie beziehen sich auf ein immer prekäres Verstehen, auf die Grenze gemeinschaftlich geteilter Wirklichkeit.<sup>42</sup>

## 2.4 Erfahrungsgestalt, Schema, Netzwerk

Der neuralgische Punkt der *CMT* ist durch den Begriff der Erfahrung bezeichnet. Ich erinnere noch einmal an die grundlegende These:

We have found that such principles are often metaphoric in nature and involve understanding one kind of experience in terms of another kind of experience.<sup>43</sup>

Das volle Gewicht der Definition hängt an der Frage, was denn überhaupt unter ‚Erfahrung‘ und ‚multidimensionaler Erfahrungsgestalt‘ zu verstehen ist. Gerade aber die Verwendung des Terminus ‚Erfahrung‘ ist in der *CMT* keineswegs eindeutig. In der angewandten Forschung zu kognitiven Metaphern ist häufig nur von „understanding one kind in terms of another kind“ die Rede – als könne man den Erfahrungsbegriff stillschweigend fallen lassen. Die erfahrungstheoretische

---

<sup>42</sup> Siehe John Dewey: *Kunst als Erfahrung*, Frankfurt/M. 1980. Eben deshalb kann das audiovisuelle Bild nicht vorderhand auf eine selbstverständlich gegebene Wirklichkeit bezogen werden, um deren Sinn – sei es im Rekurs auf semiotische, narratologische oder genrepoetische Systeme, sei es mit Blick auf die Alltagswahrnehmung strukturierende kognitive Wahrnehmungs- und Aktionsmuster – unmittelbar aus den identifizierten Repräsentationen des Bildes abzuleiten.

<sup>43</sup> Lakoff/Johnson: *Metaphors We Live By*, S. 116–117.

Begründung gerät Mal um Mal aus dem Blick – und mit ihr die genannten Bestimmungen. So ist denn häufig ungewiss, ob mit konzeptuellen Metaphern feststehende Entitäten oder operative Schemata, semantische Felder oder sinnlich konkrete Erfahrungsdomänen angesprochen werden, die man mal in kognitiven Systemen, mal in medialisierten menschlichen Interaktionen verortet.

Nicht selten tritt an die Stelle der multidimensionalen Erfahrungsgestalt das semantisch definierbare Datum einer *source domain* und einer *target domain*. Die metaphorische Interaktion gerät so zu einer rein logischen Operation, dem Mapping zweier wohldefinierter semantischer Felder.<sup>44</sup> Der Grund hierfür ist in der Verbindung von erfahrungs- und informationstheoretischen Begriffen zu suchen, die von Anfang an die Theorie kognitiver Metaphern strukturierte.

Das wiederum hat seinen Grund in der eigentümlichen Zwischenstellung, die das metaphorische Konzept selbst einnimmt. Einerseits werden metaphorische Konzepte durch ihre strukturierende Funktion innerhalb eines Modells kognitiver Informationsverarbeitung definiert; andererseits werden sie als konkrete empirische Erscheinungsformen zwischenmenschlicher Interaktion durch ihre kommunikative Funktion bestimmt. Einerseits werden metaphorische Konzepte im Interaktionsmodell eines in sich geschlossenen Systems (Organismus) mit einer (weitgehend unbestimmten) Umwelt als operatives Schema eines Kognitionsautomaten verortet; abstrakt ist dann alles, was nicht als bereits verstanden und gewusst qualifiziert werden kann. Andererseits dienen sie der intersubjektiven Strukturierung von Erfahrungen über getrennte subjektive Erfahrungshorizonte hinweg. Hier bleibt die Rede des anderen so lange abstrakt, wie sie sich mit keiner eigenen konkreten Sinneserfahrung verbinden lässt.

Faktisch sind also zwei völlig getrennte Phänomenbereiche angesprochen, wenn von metaphorischen Konzepten die Rede ist. Die Frage, die sich stellt, ist folglich, wie deren Beziehung zueinander gedacht werden soll.

Man kann das Problem recht präzise an der terminologischen Entwicklung der *CMT* nachvollziehen. Zunächst nämlich fassen Lakoff und Johnson die konkrete Konfiguration einer bestimmten Erfahrungsdomäne als eine Form erlebbarer Ganzheiten. In *Metaphors We Live By* wird die konkrete Sinnesrealität solcher Erfahrungsdomänen mit dem Begriff *multidimensional experiential gestalt* gefasst:

The multidimensional structure characterizes experiential gestalten, which are ways of organizing experiences into structured wholes.<sup>45</sup>

---

<sup>44</sup> Metaphern werden immer noch in vielen Ansätzen als rein kognitive, entkörperlichte, abstrakte Konzeptualisierungen gefasst. Siehe hierzu Gibbs' Kritik: Gibbs: Embodiment and Cognitive Science, S. 121f.

<sup>45</sup> Lakoff/Johnson: *Metaphors We Live By*, S. 81.



Die ‚multidimensionale Erfahrungsgestalt‘ benennt Phänomene, die auf der Ebene des alltäglichen Bewusstseins zugänglich und in allen Formen zwischenmenschlichen Austausches vorzufinden sind. Sie sind in ihrer perzeptiven, affektiven und kognitiven Dimension unmittelbar fassbar und können entsprechend für die Verständigung qua metaphorischer Konzepte eingesetzt werden.

Im Fortgang der Theorieentwicklung wird der Begriff durch neurowissenschaftliche Termini wie ‚sensomotorisches Schema‘ etc. verdrängt. Damit wird auch der Bezug auf die subjektive Erfahrungsrealität aufgegeben; ein Umstand, der vor allem die aktuellen Ansätze zur Metapher als *embodied meaning-making* audiovisueller Medien betrifft.

### Image schemas

Bestimmend ist in diesem Zusammenhang das Konzept der *image schemas* geworden.<sup>46</sup> *Image schemas* bezeichnen grundlegende Raum-Bewegungsbeziehungen, die als universale Grundmuster zu Bausteinen komplexer Erfahrungsmuster fungieren. Ihre Kategorisierung verdankt sich dem Bemühen, in einer Vielzahl metaphorischer Konzepte die persistenten Raum-Bewegungsmuster zu identifizieren und in einer Taxonomie zusammenzufassen. Es liegt nahe, darin ein Indiz für grundlegende Orientierungsmuster des sich in räumlichen Bezügen bewegenden Körpers zu erkennen. Die Taxonomie der *image schemas* wird in der gegenwärtigen Metaphernforschung weithin als ein grundlegendes Gerüst verkörperter Erfahrungsstrukturen in Anschlag gebracht, mit dem metaphorische Konzepte auf skelettierte Formen verkörpernder Erfahrung zurückgeführt werden.

*Image schemas* definieren mithin Schemata, die der Vielfalt „physischer Einzelerfahrungen, Sinneswahrnehmungen und motorischer Abläufe Kohärenz verleihen“;<sup>47</sup> sie adressieren eine physiologische Ebene des menschlichen

---

46 Kathrin Fahlenbrach: *Audiovisuelle Metaphern. Zur Körper- und Affektästhetik in Film und Fernsehen*, Marburg 2010; Coëgnarts/Kravanja: *Embodied Visual Meaning*; Forceville: *Metaphor in Pictures and Multimodal Representations*; Charles Forceville/Thijs Renckens: *The Good Is Light and Bad Is Dark Metaphor in Feature Films*, in: *Metaphor and the Social World* 3 (2), 2013, S. 160–179; Charles Forceville: *From Image Schema to Metaphor in Discourse: The FORCE Schemas in Animation Films*, in: *Metaphor. Embodied Cognition and Discourse*, hrsg. v. Beate Hampe, Cambridge 2017, S. 239–256.

47 Ulrike Schröder: *Kommunikationstheoretische Fragestellungen in der kognitiven Metaphernforschung. Eine Betrachtung von ihren Anfängen bis zur Gegenwart*, Tübingen 2012, S. 35 zit. n. Christina Schmitt: *Wahrnehmen, fühlen, verstehen: Metaphorisieren in der Kommunikation audiovisueller Bilder*, Dissertation, FU Berlin 2017, S. 27. Schmitts Arbeit erscheint voraussichtlich

Denkens, die allen subjektiven, sozialen oder kulturellen Differenzen, allem Verstehen und Nicht-Verstehen – kurz, die aller Erfahrungsrealität menschlicher Subjekte zuvorkommt; die folglich auch nicht im Phänomenbereich alltäglicher Erfahrungen verortet ist. Letztlich verdankt sich der Begriff einer theoretischen Operation, die in den *image schemas* die Bedingungen zu definieren sucht, die es ermöglichen, die Vielheit physischer Impulse – Sinnesreize, affektive Reflexe, Aktivierung kognitiver Schemata – überhaupt als konsistente Erfahrungsgestalt zu fassen.

In dieser Perspektive aber wird der Körper zur transzendentalen Universalie eines informationsverarbeitenden Automaten, der die Vielfalt subjektiv getrennter Erfahrungshorizonte generiert.<sup>48</sup> Im Umkehrschluss aktueller Medienanalysen werden *image schemas* zu einem System von „recurrent pattern, shape, and regularity in, or of these ongoing ordering activities“, die nun selbst wiederum zu Gegenständen von Erfahrung werden: „These patterns *emerge as meaningful structures for us* chiefly at the level of our bodily movements through space, our manipulations of objects, and our perceptual interactions.“<sup>49</sup> Das heißt nichts anderes, als dass *image schemas* im Ergebnis eine universelle Tiefensemantik („*as meaningful structures*“) bilden, aus der sich alle Sinnbildung menschlichen Denkens ableiten lässt.<sup>50</sup>

Zwar betont Johnson, dass die abstrakten, tiefliegenden Strukturen lediglich kohärenzstiftende Operationsmuster darstellen, die allein deshalb keine fixen Gegebenheiten sein könnten, weil sie höchst unterschiedliche Erfahrungen strukturierten; sie seien vielmehr als flexible Struktur zu denken, die in je gegebenen Konstellationen anderen situationsgebundenen Modulationen unterliege. Erst die prozessuale Anpassung an ständig wechselnde Situationen mache die *image schemas* als ein generatives System denkbar.

Aber die Grenze zwischen ‚kinästhetischen‘ Verlaufsmustern (Orientierung, Bewegung, Balance in der Selbstwahrnehmung eines sich im Raum bewegenden Körpers), kognitiven Operationen und sinnhaften Gestalteinheiten wird systematisch verwischt, wenn mit Blick auf solche Muster von „primary means“<sup>51</sup> die

---

2019 mit dem Titel *Wahrnehmen, fühlen, verstehen: Metaphorisieren und audiovisuelle Bilder* bei de Gruyter.

48 Christina Schmitt hat differenziert dargelegt, wie aus den Versuchen, die metaphorische Bedeutungskonstruktion auf reine Verkörperungsprozesse zurückzuführen, im Ergebnis eine universelle Tiefensemantik hervorgeht. Vgl. Schmitt: *Wahrnehmen, fühlen, verstehen*, S. 20–31.

49 Mark Johnson: *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason*, Chicago 1987, S. 31 [meine Hervorhebung, H. K.].

50 Schmitt: *Wahrnehmen, fühlen, verstehen*, S. 28.

51 Johnson: *The Body in the Mind*, S. 30.

Rede ist. Die Stabilität eines solchen Systems von *primary means* kann dann doch wieder nur als Konventionalisierung gedacht und begründet werden: „They are not eternally fixed objects, as Objectivism suggests, but they gain a certain stability by becoming conventionally located in our network of meaning.“<sup>52</sup> Auch für Johnson ist also klar, dass *image schemas* nur dann die Ebene verkörpernder kognitiver Prozesse beschreiben können, wenn sie als grundlegende operative Muster verstanden werden.

Man darf von der experimentellen Forschung durchaus erwarten, dass sie Schemata (der Orientierung, Bewegung und Balance im Raum) beschreiben kann, die menschliche Wahrnehmungs- und also auch Kognitionsprozesse strukturieren. Doch werden sich diese nicht als *primary means* zu einer Tiefensemantik systematisieren lassen. Liegen doch alle semantischen Effekte *per definitionem* auf der Ebene des Alltagsbewusstseins.

Noch die taxonomisch erfassten Figurationen von Selbst-Bewegung und Bewegung folgen einem konventionellen Raum-Zeit-Schema, das sich keineswegs unmittelbar auf die Selbstwahrnehmung eines im Raum sich bewegenden Ich zurückführen lässt. Eingeschaltet ist immer das Selbst der Selbstwahrnehmung, d. i. das Monitoring der affektiven und perzeptiven Reize und deren subjektive Interpretation als ‚mein Körper in dieser Situation‘. Die Interpretation stützt sich auf konventionalisierte Beschreibungssysteme – wie immer auch strukturiert, gerahmt oder gemodelt von kognitiven operativen Mustern. Von den Konventionalisierungen ist die Raum-Zeit-Matrix, in der Bewegung als Verlagerung von Körpern in einem homogenen und unveränderbaren Raum auf der Zeitachse gemessen wird, sicher eine der grundlegendsten; sie ist deshalb aber nicht weniger eine Konvention, auf die sich die subjektive Erfahrung des eigenen Körpers bezieht. Nicht zuletzt die experimentelle neurowissenschaftliche Forschung ist auf völlig andere Raum-Zeit-Modelle angewiesen, um neuronale Prozesse beschreiben zu können (z. B. als Intensitätswechsel neuronaler Aktivitäten).<sup>53</sup>

Wenn *image schemas* tatsächlich als bedeutungsgenerierende Schemata verstanden werden, verweisen sie also auf die Ebene des Monitorings der affektiven Impulse und Sinnesdaten. Das aber ist die Ebene eines Bewusstseins vom eigenen

---

52 Johnson: *The Body in the Mind*, S. 30.

53 Vgl. Shahar Arzy/Gregor Thut/Christine Mohr/Christoph M. Michel/Olaf Blanke: *Neural Basis of Embodiment: Distinct Contributions of Temporoparietal Junction and Extrastriate Body Area*, in: *The Journal of Neuroscience* 26 (31), 2006, S. 8074–8081; Mirta Villarreal/Esteban A. Fridman/Alejandra Amengual/German Falasco/Eliana Roldan Gerschcovich/Erlinda R. Ulloa/Ramon C. Leiguarda: *The Neural Substrate of Gesture Recognition*, in: *Neuropsychologia* 46 (9), 2008, S. 2371–2382; Paula M. Niedenthal: *Embodying Emotion*, in: *Science* 316 (5827), 2007, S. 1002–1005.

Körper als eines ‚Körpers, der ich selber bin‘ („Leibbewußtsein“); die Ebene eines subjektiven, von allen anderen unterschiedenen Erfahrungshorizonts.

### Netzwerke

Genau in dem Maße, in dem die physiologischen Aspekte des Embodiment ins Zentrum der Anwendung der CMT rücken, tritt erneut die Erfahrungsdimension in den Vordergrund, d. i. das sich in seinem eigenen körperlichen In-der-Welt-Sein gewahr werdende Subjekt.<sup>54</sup> Statt selbstverständliche Daten zu decodieren, entfaltet die metaphorische Operation notwendig ästhetische Korrelationen, d. h. affektive und perzeptive Ähnlichkeits- und Vergleichsrelationen, die sich nicht auf der Ebene rein logischer Bedeutungsrelationen fassen lassen: Metaphern überblenden zwei Wahrnehmungsfigurationen, die in ihrer affektiven, perzeptiven und sinnhaften Dimension als konkrete, verkörperte Erfahrungen aufeinander bezogen werden.

Mit ihrem Konzept der „audiovisuellen Metaphern“ hat Kathrin Fahlenbrach einen Ansatz entwickelt, der von der Konfiguration perzeptiver, affektiver und kognitiver Dimensionen der Wahrnehmung ausgeht. Ihre Arbeit gehört sicher zu den differenziertesten Versuchen, die CMT für die Analyse audiovisueller Medien fruchtbar zu machen. Entgeht sie doch der Versuchung, neurowissenschaftlich inspirierte Konzepte des Embodiment pauschal als Begründung eines Modells verkörpernder Rezeption audiovisueller Bilder ins Feld zu führen.<sup>55</sup>

Fahlenbrach stützt sich auf ein psychologisches Wahrnehmungsmodell, das Verstehensprozesse als Interaktion innerhalb eines Netzes kognitiver (d. h. verstandeslogischer), perzeptiver und affektiver Operationen begreift. Konzeptuelle Metaphern sind dann als kognitive Mechanismen (festgelegte Operationsstrukturen des Mapping) zu verstehen, die zwei Netzwerke miteinander korrelieren; sie ‚mappen‘ ein bestimmtes Netz kognitiver, affektiver und perzeptiver Regelprozesse mit dem Netz eines anderen Erfahrungsbereiches. Die konzeptuelle Metapher basiert also keineswegs auf einer rein verstandesmäßigen Operation; sie umfasst vielmehr die Interaktion zwischen Erfahrungsbereichen in ihren

---

<sup>54</sup> Gibbs hat das sehr eindrücklich in seiner Hinführung zum Konzept des Embodiment dargestellt. Vgl. Gibbs: *Embodiment and Cognitive Science*; Mark Johnson: *The Meaning of the Body. Aesthetics of Human Understanding*, Chicago 2007.

<sup>55</sup> Wie das etwa bei Coëgnarts/Kravanja der Fall ist, wenn sie Spiegelneuronen ins Feld führen, um zu erklären, wie Zuschauer *image schemas* realisieren, die ein Regisseur intuitiv nutze, um seine Ideen in audiovisuellen Bildern zu vermitteln. Vgl. Coëgnarts/Kravanja: *Embodied Visual Meaning*; Coëgnarts/Kravanja: *Embodied Visual Meaning in Film*.

sinnlichen, affektiven und semantischen Dimensionen. Damit ist recht präzise die ‚multidimensionale Erfahrungsgestalt‘ definiert, von der Lakoff und Johnson sprechen.

Wie diese begreift auch Fahlenbrach konzeptuelle Metaphern als ein kognitives System, das als apriorische Struktur fungiert, welche die metaphorische Bedeutungskonstruktion in konkreten Kommunikationen überhaupt erst ermöglicht. Der Widerspruch zwischen informationstheoretischem (kognitives System) und erfahrungstheoretischem Konzept wird nun explizit hin zu einer „körperlichen Tiefensemantik“ aufgelöst. Primärmetaphern<sup>56</sup> und kognitive Schemata sind die Bausteine eines *primary meanings*, auf das sich komplexe konzeptuelle Metaphern gründen.

Audiovisuelle Bilder wiederum weisen „mit ihren abbildhaften Bildern und Klängen“<sup>57</sup> eine große Affinität zu den körperbasierten Wahrnehmungskonfigurationen konzeptueller Metaphern auf. Die alte Idee der Metapher als Bild oder bildhafter Ausdruck ist hier überführt in die Gleichung von audiovisuellem Bild und konzeptueller Metapher, die beide eine Konfiguration kognitiver, perzeptiver und affektiver Schemata darstellen. Audiovisuelle Metaphern lassen sich so als Netzwerke perzeptiver, affektiver und kognitiver Operationsmuster definieren, in denen konzeptuelle Metaphern repräsentiert und reflektiert werden.

Konzeptuelle Metaphern können demnach – in abgestuften Graden der Komplexität – als apriorische kognitive Struktur vorausgesetzt werden, die Medienproduzenten kalkuliert einsetzen, um Zuschauer in ihrer unmittelbaren körperlichen Reaktion in die Narration (das meint hier immer die repräsentierte Handlung oder den repräsentierten Sachverhalt) eines Films oder Videos einzubeziehen. Die Zuschauer antworten genau mit jenen unwillkürlichen und ‚reflexhaften‘ Vernetzungen kognitiver, affektiver und perzeptiver Abläufe, die in der audiovisuellen Metapher repräsentiert sind. D. h., audiovisuelle Metaphern aktivieren „die universalen Muster des körperlichen und affektiven Erlebens“, die „kognitiv verankerte konzeptuelle Metaphern [...] beinhalten.“<sup>58</sup>

Die audiovisuelle Metapher wird letztlich als Funktionale definiert, die dem Zweck dient, das repräsentierte Narrativ (die abstrakte Idee des Filmemachers oder die Handlung) zu vermitteln. Sie fungiert als *audiovisuelle Konkretion abstrakter Ideen von Medienmachern* und sorgt für die affektive Einbindung der Zuschauer in die mediale Darstellung. Letztlich wird damit das traditionelle Sender-Empfänger-Schema auf ein wirkungsästhetisches Kommunikationsmodell

---

<sup>56</sup> Zum Begriff der *primary metaphor* vgl. die Fußnote 93 im dritten Teil dieses Buches.

<sup>57</sup> Fahlenbrach: Audiovisuelle Metaphern, S. 45.

<sup>58</sup> Fahlenbrach, S. 11 f.

übertragen: das Kalkül der Medienproduzenten, die über intuitives Wissen resp. rhetorische Konvention verfügen, und die Körper von Rezipienten, die einem universellen wirkungsästhetischen Schematismus folgen.

### Die Metapher des Kognitionsautomaten

Es ist eine Sache, die Hypothese zu vertreten, dass *image schemas* eine basale Aktivitätsstruktur menschlicher Kognition bezeichnen, die am Grund aller Metaphern wirksam ist; eine ganz andere Sache ist es, im Verweis auf diese Hypothese die Bedeutung einer audiovisuellen Konfiguration feststellen zu wollen. Es ist eine Sache, zu versuchen, eine solche Hypothese durch deskriptive Verweise zu erhärten, indem ich die prinzipielle Rückführbarkeit auf *image schemas* an einer Vielzahl metaphorischer Ausdrücke durchspiele; eine ganz andere Sache ist es, die Bedeutung oder die Funktionsweise konkreter audiovisueller Darstellungen zu erklären, indem das System kognitiver Schemata als generative Tiefensemantik in Anschlag gebracht wird.

Auch die Theorie der „audiovisuellen Metaphern“ reproduziert damit die entscheidende Schwachstelle der *CMT*: Sie folgt eben jenem Schematismus strukturalistischer Sprachtheorien, gegen die sich die *CMT* ursprünglich gewendet hat. Letztlich spielt es nämlich keine Rolle, ob ich eine Transformationsgrammatik oder eine Tiefensemantik als ‚inneres Wissenssystem‘ veranschlage; so oder so folge ich dem Schema repräsentationaler Denkvorstellungen, mit denen die Welt gespalten wird „in eine ‚Tiefenstruktur‘, die ein universelles Muster birgt und eine ‚Oberfläche‘, die dieses Muster unter jeweils konkreten [...] Umständen aktualisiert.“<sup>59</sup> Selbst wenn man dem Umstand, dass die empirischen Befunde neurowissenschaftlicher Forschung keineswegs dazu berechtigen, von einem statischen System primären Bedeutens auszugehen, kein allzu großes Gewicht bemisst – die ‚Spaltung der Welt‘ in eine generative und eine phänomenale Struktur hat schwerwiegende Konsequenzen, nicht nur, aber doch in besonderem Maße für das Verständnis des Verhältnisses von audiovisuellem Bild und Metapher.

Letztlich wird damit ein Kommunikationsmodell perpetuiert, das die Akteure als in sich geschlossene, solipsistische Kognitionsautomaten entwirft, auch wenn sie mit einem empfindsamen Körper ausgestattet werden. Der Körper nämlich fungiert nun selbst in seinen unwillkürlichen affektiven Reaktionen und kogni-

---

<sup>59</sup> Sybille Krämer: Sprache – Stimme – Schrift: Sieben Gedanken über Performativität als Medialität, in: Performanz. Zwischen Sprachphilosophie und Kulturwissenschaften, hrsg. v. Uwe Wirth, Frankfurt/M. 2002, S. 323–346, hier S. 324.

tiven Mustern als Decodierungsmaschine (die unwillkürliche Aktivierung kognitiver Schemata auf Seiten der Empfänger), die Medienmacher (Sender) intuitiv und kalkuliert als Codierungssystem nutzen.<sup>60</sup> Zweifellos handelt es sich hierbei um eine höchst abstrakte Vorstellung von Körper; sie versteht den Körper als eine bedeutungsgenerierende Matrix, die letztlich als apriorische Bestimmung sinnvoller Kommunikation fungiert. Tatsächlich wäre im strengen Sinne von einem transzendentalen Körper zu sprechen, der lediglich als Natur vorgestellt wird.

Wenn man *image schemas* als ein System von *primary means* auffasst, sind sie nichts anderes als ein von Forschern genutztes arbiträres Bezeichnungssystem. Die in solchen Untersuchungen gebräuchliche Taxonomie schließt genau in dem Maße Konventionalisierungen ein, wie sie als tiefensemantisches System für die Interpretation audiovisueller Repräsentationen in Anschlag gebracht wird. Deshalb der notorische Verweis auf die Konvention, ohne den kognitionstheoretische Analysen sinnkonstituierender Prozesse in audiovisuellen Medien nicht auskommen. Die Prozesse der Konventionalisierung freilich verschwinden in der Blackbox eines Embodimentmodells, das den Körper als informationsverarbeitenden Automaten definiert, der einer mehr oder weniger unterbestimmten Umwelt gegenübersteht. Mit Blumenberg könnte man sagen, dass die CMT mit anderen Kognitionswissenschaften eine erkenntnisstrukturierende Metapher teilt, die den immer gleichen blinden Fleck entstehen lässt: der Rechner, der Computer, der informationsverarbeitende Automat ist das grundlegende Modell menschlicher Kognition.

In dieser Perspektive lassen sich alle historischen und kulturellen Ausformungen metaphorischen Denkens, und damit auch Bewegungsbild-Metaphern, den allgemeingültigen Schemata menschlicher Kognition subsumieren – und paradoxerweise zugleich durch diese erklären. Deren ‚eigentliche‘ Bedeutung lässt sich stets im Rückgang der Analyse auf universale Schemata kognitiver Prozesse mit größter Eindeutigkeit feststellen. Ganz gleich, welchen Sinn die Medienrezipienten im Gebrauch der Metaphern mit diesen herstellen mögen. Ihr *Machen* wird wie alle historisch-kulturelle Sinnproduktion als zu vernachlässigende Größe systematisch aus der Betrachtung ausgeschlossen (eine notwendige Strategie – so könnte man mit Blick auf de Certeau anfügen –, wenn man die Beherrschbarkeit der Sinnproduktion durch das Kalkül von Medienmachern und die Methoden von Wissenschaftlern beschreiben und festschreiben will). Der historische Tie-

---

<sup>60</sup> Vgl. Forceville: Visual and Multimodal Metaphor in Film; Maarten Coëgnarts/Peter Kravanja: Towards an Embodied Poetics of Cinema: The Metaphoric Construction of Abstract Meaning in Film, in: Alphaville. Journal of Film and Screen Media (4), 2012, <http://www.alphavillejournal.com/Issue%204/PDFs/ArticleCoegnarts&Kravanja.pdf> (31. März 2018).

fenraum kultureller Sinnproduktion wird mit Verweis auf ein allseits verfügbares kulturelles Genre-, Medien- oder Alltagswissen zur akzidentiellen Restmenge, die in ihrer systematischen Unterbestimmtheit jeder noch so löcherigen Analyse audiovisueller Metaphern als komplementäre Sinnfülle zur Verfügung steht.

Ausgeblendet wird schließlich auch jene Relation, die Fahlenbrach ins Zentrum ihres Modells konzeptueller Metaphern gerückt hat: die Kongruenz zwischen konzeptueller Metapher und audiovisuellem Bild. Ist doch eine solche Kongruenz nur auf der Ebene der Erfahrung des konkreten Wahrnehmungslebens vorstellbar – nicht aber auf der Ebene eines kognitiven Systems.

### **Zur Performativität des Bewegungsbildes**

Das audiovisuelle Bewegtbild ist in seiner Medialität selbst noch als eine Konfiguration von Perzepten, Affekten und Konzepten zu verstehen, die nicht erst über abbildliche Repräsentationen vermittelt werden. Eine Konfiguration des Sehens und Hörens, des Fühlens und Denkens, die notwendig von Medienkonsumenten zu realisieren ist, sofern sie das audiovisuelle Bild überhaupt rezipieren, d. h. wahrnehmen. Die Kongruenz von audiovisuellem Bild und konzeptueller Metapher ist weniger als Nähe, denn als tatsächliche Strukturgleichheit zu fassen, die auf ein und derselben Ebene zu verorten ist: Beide nämlich sind als perzeptive, affektive und kognitive Konfigurationen eines spezifischen Wahrnehmungsszenarios zu begreifen (als ‚multidimensionale Erfahrungsgestalt‘, sprich audiovisueller Bildraum).

In diesem Sinne spricht Dorothea Horst von „audio-visual figurativity“.<sup>61</sup> Sie benennt damit zum einen eine Beobachtungsebene, in der bildliche und sprachliche Ausdrücke aufeinander beziehbar sind. (Der Audiovisualität des Fernsehbildes wird die figurative Sprachbildung analog gesetzt.) Vor allem aber fasst sie damit das Verhältnis von audiovisuellen Bewegtbildern und Repräsentationen selbst als eine mediale Konfiguration: Denn es ist nicht so, dass die repräsentierten Gegenstände audiovisueller Bilder bereits als solche – eben als für sich gegebene, aneinander gefügte Bild-Repräsentationen – figurative Ausdruckskonstellationen abbilden. Vielmehr sind die audiovisuellen Bewegtbilder in ihrer spezifischen medialen Performanz, in der Inszenierung des audiovisuellen Bildes, als eine ‚audiovisuelle Figuration‘ anzusprechen. Sie sind den Rezipienten unmittelbar als Sinneskonfigurationen, d. h. als multidimensionale Erfahrungsgestalt

---

<sup>61</sup> Vgl. Dorothea Horst: *Meaning-Making and Political Campaign Advertising. A Cognitive-linguistic and Film-analytical Perspective on Audiovisual Figurativity*, Berlin/Boston 2018.



körperlicher Resonanzen zugänglich, während jeder Versuch, solche Resonanzen im Rückgriff auf zugrundeliegende kognitive Systeme zu erklären, immer in die mehr oder weniger gut begründete spekulative Introspektion anderer Intentionalitäten führt<sup>62</sup> – was auch nur die komplizierte Umschreibung eines unreflektierten hermeneutischen Verfahrens ist.

Ein System generativer Primärstrukturen, mit denen sich die Bedeutung metaphorischer Konzepte feststellen lässt,<sup>63</sup> ist eine theoretische Fiktion, die zumindest bezogen auf audiovisuelle Bilder zu einer inadäquaten Gegenstandsbestimmung führt. Sie übergeht nämlich die entscheidende Dimension ihrer Medialität: die Bewegung des Bewegungsbildes, seine Temporalität.

Christina Schmitt hat mit Blick auf die konzeptuelle Metapher das skizzierte Problem als Gegensatz von kognitivem System und prozessualer Performanz diskutiert. Sie bezieht sich dabei auf Sybille Krämer, die von der „ästhetischen Performativität“, der Oberflächenerscheinung medialer Kommunikation spricht.<sup>64</sup> Das meint die mediale Modulation aller Formen der Kommunikation, die Ausdrucksqualitäten medialer Oberflächen. Audiovisuelle Bilder können dann als ein „in-Szene-setzendes Wahrnehmbarmachen“ verstanden werden; dabei meint die Performativität der Medien einen Akt des „Phänomenalisierens“,<sup>65</sup> d. h., sie ‚machen wahrnehmbar‘, indem sie Verkörperungsprozesse in Gang setzen. Bedeutung entsteht also nicht im Rückgang auf ein drittes gemeinsames System (Sprache, kognitive Schemata, *primary meanings*); „sie ist keine mehr oder weniger stabile Entität, sondern existiert nur in der flüchtigen und prozessualen Gegenwärtigkeit des Medienumgangs.“<sup>66</sup>

Will man die *CMT* als Analyseverfahren auf audiovisuelle Bilder beziehen, ist deshalb zuallererst das audiovisuelle Bild selbst als ein spezifisches „in-Szene-setzendes Wahrnehmbarmachen“ theoretisch zu fassen. Ohne einen medientheoretisch fundierten Begriff vom audiovisuellen Bewegtbild verschwindet jede

---

**62** Neurophilosoph Stephan Schlemm sagt, alle bewussten Wahrnehmungserlebnisse „sind erst einmal nur der Person oder dem Lebewesen zugänglich, die oder das sie hat.“ (zit. n. Saskia Gerhard: „Vergesst den freien Willen!“, in: Zeit Online, 18. Oktober 2017, <http://www.zeit.de/wissen/2017-10/was-ist-bewusstsein/komplettansicht> (31. März 2018)). Auf die Grenzen des Bewusstseins der Anderen stößt auch der Zuschauer, der Macher etc. Gerade weil es immer verkörpertes Bewusstsein, Bewusstsein eines eigenen Körpers ist, erfordert es ein solches Verstehen.

**63** Quasi als Codesystem, das die Analyse decodiert.

**64** Vgl. Schmitt: Wahrnehmen, fühlen, verstehen, S. 84–87.

**65** Sybille Krämer: Was haben „Performativität“ und „Medialität“ miteinander zu tun? Plädoyer für eine in der „Ästhetisierung“ gründende Konzeption des Performativen, in: Performativität und Medialität, hrsg. v. Sybille Krämer, München 2004, S. 13–32, hier S. 25.

**66** Krämer, S. 25.

mediale und modale Differenz. Mehr noch: Es verschwindet die basale Struktur, die das audiovisuelle Bild als ein in sich bewegtes Bild zuallererst begründet. Insofern gilt für die angewandte CMT im Besonderen, was im ersten Teil dieses Buches über die kognitive Filmtheorie im Allgemeinen gesagt wurde: Ihr gerät im kognitiven Schematismus aus dem Blick, was doch eigentlich untersucht werden soll: das audiovisuelle Bild als ein *Bewegungsbild*.

Quer zu den unterschiedlichen medialen Formaten, vom Webcam-Video bis zum Kinofilm, zeichnen sich audiovisuelle Bilder gerade dadurch aus, dass sie keine Sukzession isolierter, stillgestellter Bild-Repräsentationen sind, sondern in der permanenten Modulation ihrer Ausdrucksformen multiple Raum-Zeit-Konfigurationen hervorbringen. Die Bewegung audiovisueller Bewegungsbilder impliziert immer schon die dynamisch-energetische Veränderbarkeit solcher Raum-Zeit-Figurationen. Darin liegt ihre spezifische Temporalität begründet. Diese Zeitlichkeit der Bewegung ist konstitutiv für den medialen Modus ihrer Repräsentationen. Entsprechend muss sich jeder Versuch, die Repräsentationen audiovisueller Bilder zu fassen, zunächst auf die Rekonstruktion der raum-zeitlichen Konfiguration in ihren je spezifischen Modulationen als Bewegungsbild beziehen, um überhaupt bestimmen zu können, was denn letztlich Gegenstand der Repräsentation ist.

Wie im ersten Teil des Buches erläutert, unterscheiden wir deshalb begrifflich zwischen *kinematografischem* oder *audiovisuellem Bewegungsbild* (*moving image*) und *filmischem Bewegungsbild* (*Movement-Image*).<sup>67</sup> Mit dem Bewegungsbild sind die medialen Artefakte bezeichnet, d. h. alle audiovisuellen Bilder, ganz gleich welchen medientechnischen Typus; durchaus in dem Sinne, dass sich diese Bilder als isolierte abbildliche Repräsentationen in sich bewegter Sachverhalte verstehen lassen. Der Begriff des Bewegungsbildes hingegen bezeichnet den übergreifenden Zusammenhang der zeitlichen Gestalt eines Films, Videos oder einer Fernsehreportage; er verweist auf das inszenatorische, poetische oder rhetorische Regime, unter dem die isolierten Bewegungsbilder – d. h. die audiovisuellen Repräsentationen – als Elemente eines in der Zeit der Rezeption sich entfaltenden Ganzen eines filmischen Bewegungsbildes erfasst werden.

### **Metamorphosen des Bildraums: DEATH AND THE MOTHER**

Im Folgenden möchte ich das an einer weiteren analytische Skizze erläutern: Gegenstand der Betrachtung ist Ruth Lingfords kurzer Animationsfilm *DEATH*

---

<sup>67</sup> Den Begriff entlehnen wir Deleuze. Vgl. Deleuze: Das Bewegungs-Bild.

AND THE MOTHER (GB 1997).<sup>68</sup> Ich wähle dieses Beispiel zum einen, weil der Animationsfilm in besonderer Weise die Differenz zwischen der abbildlich repräsentierten Bewegung – der Verlagerung von Körpern im gegebenen Raum und deren handlungslogische Interpretation – und der Dynamik des Bewegungsbilds selbst zum Thema hat. Zum anderen ist Lingfords Video Gegenstand einer Analyse, die sich auf die Taxonomie der *image schemas* als bedeutungsgenerierendes System stützt.<sup>69</sup> Im Vergleich sollte also der Unterschied, der sich aus einer erfahrungstheoretischen gegenüber einer informationstheoretischen Lesart der *CMT* ergibt, klar hervortreten.

DEATH AND THE MOTHER beruht auf Zeichnungen, die das künstlerische Verfahren des Holzschnitts nachahmen. Dieses Nachahmen ist im Hinblick auf Poetik und Ästhetik des kurzen Videos prägend: Die audiovisuellen Bilder gründen hier auf einem höchst abstrakten, dynamischen Wechselspiel zwischen hellen und dunklen Flächen, dichten und weniger dichten Schraffuren. Doch ist es vor allem der beständig wechselnde Status des Verhältnisses von bloßer Auslassung (*negative space*) und gestalthafter Figur, Schraffur und Figuration innerhalb der audiovisuellen Bewegung, von Positiv und Negativ des Bildes – die Leere der Auslassung wird im nächsten Moment zur gestalthaften Fülle einer abbildlichen Figuration und umgekehrt – bzw. das Prinzip dieses Wechsels selbst, was die Bedeutung einzelner audiovisueller Bilder, d. h. instantaner Wahrnehmungsgestalten, innerhalb der poetischen Logik des filmischen Bewegungsbildes prägt. Erst die animierte Bewegung lässt aus den Schraffuren und Schattierungen gegenständliche Figurationen entstehen, die nun ihrerseits in einer permanenten Metamorphose – und in der übergreifenden Bewegung jener Metamorphosen – sich fortwährend verändernde Haltungen und affektive Dynamiken zum Ausdruck bringen.

Wir, die Zuschauer, erkennen mühelos die Figurationen einer märchenhaften Handlung: das Ringen zwischen Mutter und Tod um das Leben des Kindes. Dies ist jedoch nur deshalb der Fall, weil wir ein diffuses Wissen um allegorische Figuren und Ikonografien aus Märchen und Legenden einbringen, während wir zugleich in jedem Moment affektiv in die filmischen Gesten und Gebärden eines archetypischen Beziehungsmusters verwickelt sind. Erst in dieser Konstellation, d. h. in der Verbindung von kulturhistorischem Wissen und gegenwärtiger ästhetischer Erfahrung, kann die animierte Bewegung des audiovisuellen Bildes, das sich dynamisch wandelnde Verhältnis von Mustern, Schraffuren und Leerstellen zur Ikonografie einer Mutter werden, die ihr sterbendes Kind in den Armen hält

<sup>68</sup> Online abrufbar unter <https://www.youtube.com/watch?v=hi4zdX-OG5Q> (02. April 2018).

<sup>69</sup> Vgl. Forceville: From Image Schema to Metaphor in Discourse.

und sich dagegen wehrt, es dem Tod zu überlassen. Die Figuration entsteht aus einer radikal dynamisierten Schwarz/Weiß-Flächenstruktur im Stil des Holzschnittdrucks.

Nicht in den einzelnen Repräsentationen, sondern in eben jener Dynamik des ständigen Wandels, der Kippfiguren und Metamorphosen, liegt das Affizierungspotential der audiovisuellen Bilder begründet.<sup>70</sup> Ist doch in jedem Zug das Erscheinen der Figuration mit der dynamisierten Wahrnehmungsempfindung eines Rezipienten verknüpft, der unentwegt als Figur zu fassen sucht, was im nächsten Moment nurmehr Schraffur und Fläche ist – und vice versa. D. h., es gibt notwendig eine permanente Verknüpfung zwischen den Metamorphosen der animierten Holzschnittmuster und der Dynamik sich wandelnder Wahrnehmungsempfindungen des Rezipienten. Das Bewegungsbild geht bereits aus einer Interaktion von audiovisuellen Bewegtbildern (d. h. der Animationen) und Wahrnehmungsprozessen (d. h. der fließenden Metamorphose der einzelnen Animationsphasen innerhalb der Wahrnehmung des Zuschauers) hervor; die affizierende Kraft der technisch animierten Bewegung und die Dynamik des rezeptiven Prozesses als affizierter Körper sind untrennbar ineinander verwoben.

Setzt nun aber die Analyse, gewissermaßen in Umkehrung des Wahrnehmungsprozesses, bereits auf der Ebene der repräsentierten Handlung des Märchens an – eine Mutter, die ihr Kind dem Tod überlassen soll –, so lässt sie gerade den kognitiven Prozess außer Acht, der aus dem Schwarz-Weiß-Schema von hellen und dunklen Schraffuren ein Bewegungsbild entstehen lässt. Erst auf dem Grund der Syntheseleistung der Rezeption können sich die figurativen, gestischen Ausdrucksbewegungen<sup>71</sup> einer Mutter, die um ihr Kind ringt, abheben. Will man in der repräsentierten Handlung Bewegungsmuster identifizieren, die der Taxonomie der *image schemas* entsprechen, verfehlt man Raum, Zeit und Bewegung des audiovisuellen Bildes.<sup>72</sup>

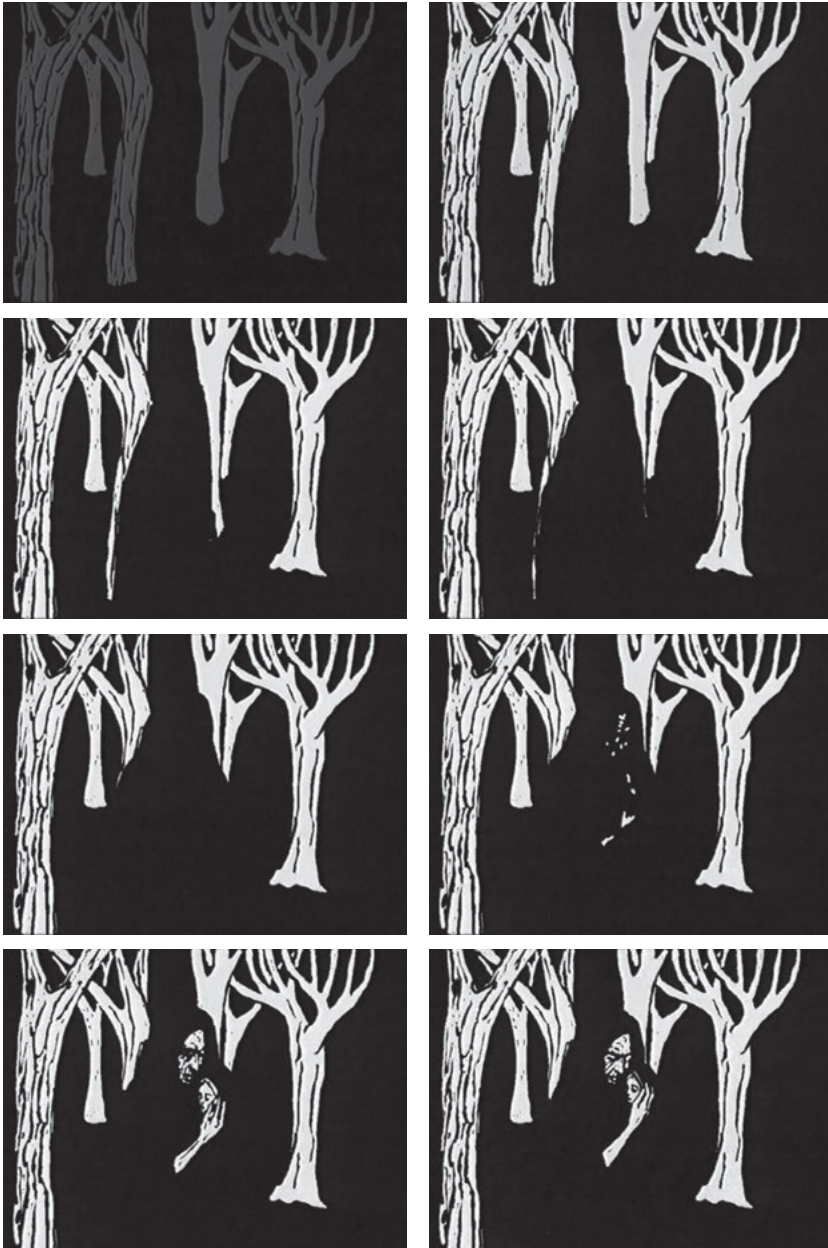
Anders gesagt: Einer Analyse, die sich unmittelbar auf die repräsentierte Handlung von *DEATH AND THE MOTHER* bezieht, entgeht nicht nur der Rekurs auf die alte Holzschnitttechnik als Bildtypus, der im Schwarz-Weiß-Schema aufgerufen

---

<sup>70</sup> Vgl. McLuhan zum Figur-Grundverhältnis: Marshall McLuhan: Das resonierende Intervall, in: absolute Marshall McLuhan, hrsg. v. Martin Baltes, Rainer Höltzschl, Philip Marchand, Freiburg 2002, S. 210–218; McLuhan fasst die Wirkung des Mediums, sein Affizierungspotential, als den Grund, auf dem die Figur erscheint.

<sup>71</sup> Zum Konzept der Ausdrucksbewegung vgl. die Seiten 133–143.

<sup>72</sup> Vgl. Forceville: From Image Schema to Metaphor in Discourse, S. 245–247.



**Abb. 13:** DEATH AND THE MOTHER

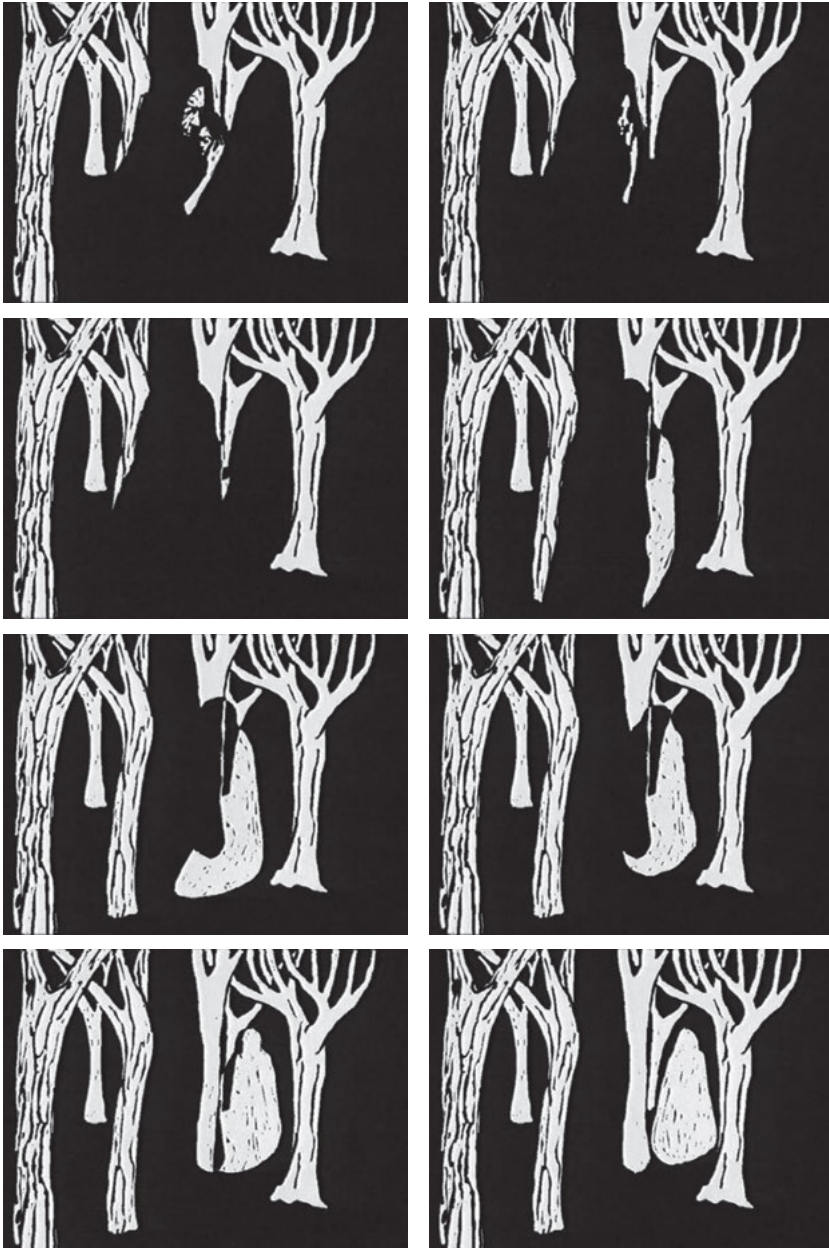


Abb. 13: DEATH AND THE MOTHER

ist; ihr entgeht nicht nur der Reflexionsraum, der sich eröffnet, wenn man auf dieser Linie die lange Geschichte der Poetik grafischer Druckkunst auf aktuelle Verfahren der Medienkunst bezieht. Nein, ihr entgehen auch die vielfältigen bildhistorischen Assoziationen, die notwendig in den Prozess einfließen, der aus der Komposition dynamisch sich verändernder Holzschnittmuster – welche beispielsweise an die Ikonografie zahlloser Pietà-Darstellungen erinnern – das Bild einer leidenden Mutter werden lässt. Dabei ist es zunächst egal, ob diese Ikonografie eine philologisch reflektierte Linie abendländischer Tradition bis zu den Varianten des Persephone-Mythos zieht, oder ganz in diffuser Allusion verbleibt. Entscheidend ist jedoch der Umstand, dass die kulturhistorische Dimension die Parameter der Sinnbildung selbst beeinflusst.

Auf all das mag eine kognitive Metaphernanalyse mit dem Verweis auf Konvention und Wissen von Machern und Rezipienten verzichten können. Das wird man jedoch kaum von diesem letzten Punkt behaupten können: *Einer Analyse, die sich unmittelbar auf isolierte Repräsentationen bezieht, entgeht die affizierende Wirkung, die sich in der Interaktion von animierten Metamorphosen und rezipierendem Empfindungskörper entfaltet.*

Gerade mit Blick auf die Affizierung der Zuschauer wird die Differenz zwischen informationstheoretischem und erfahrungstheoretischem Analysemodell deutlich: Was man – aus kognitionstheoretischer Perspektive – den allegorischen Figuren von *DEATH AND THE MOTHER* erst *im Nachgang* zuschreiben mag (man findet die repräsentierte Szene anrührend, bewegend, traurig), ist als ein Gefühl für das Bewegungsbild in seiner Gänze von Anfang an eine Dimension der Dynamik des Wahrnehmungsprozesses. Der Affekt geht unmittelbar aus der Interaktion von animierter Bewegung und Wahrnehmungsempfindung, aus der Koproduktion von audiovisuellem Bild und Zuschauerkörper hervor. Die Dynamik dieser Interaktion in der Dauer der Entfaltung des audiovisuellen Bildes bildet buchstäblich den Grund, auf dem die audiovisuelle Figuration als märchenhaft anrührende Allegorie erst wahrnehmbar wird.

### **Der Modus des Bewegungsbildes**

Das Bewegungsbild ist immer auf in sich veränderliche Raum-Zeit-Konstruktion bezogen – d. h., es bezieht sich auf Bewegung als dynamisch-energetische Veränderbarkeit von Raum-Zeit-Konfigurationen. Die Metamorphose des Bildraums bildet die ästhetische und poetologische Rahmung. Sie umfasst die einzelnen audiovisuellen Repräsentationen, zieht sie hinein in einen übergreifenden Wahr-

nehmungszusammenhang und lässt sie zu Ausdrucks- und Bedeutungsträgern der je spezifischen kompositorischen oder poetischen Ordnung werden. Deshalb auch kann sich die Analyse audiovisueller Bilder nicht unmittelbar auf textuelle, narratologische, semiotische oder ikonografische Systeme stützen. Sie muss den Umweg über die Rekonstruktion des Prozesses ästhetischer Erfahrung von Zuschauern gehen.

Für die Analyse audiovisueller Bilder ist ein Begriff vom Bewegungsbild in Anschlag zu bringen, der die unterschiedlichen Modalitäten (Ton, Geräusch, Sprache, Bewegungsfigurationen, visuelle Muster der Alltagswahrnehmung, Ikonografien) nicht isoliert ausmacht, um sie anschließend in ihren wechselnden Bezüglichkeiten zu beschreiben; vielmehr gilt es, die übergreifenden Dynamiken zu adressieren, welche die Modalitäten in einem audiovisuellen Wahrnehmungsganzen, einer ‚multidimensionalen Erfahrungsgestalt‘ aufgehen lassen. Kurz: Es bedarf eines analytischen Zugangs, der audiovisuelle Bilder nicht als additives Konglomerat unterschiedlicher Medialitäten (wie Text, Sprache oder fotografische Abbildlichkeit) begreift, sondern deren spezifische Temporalität als basale Eigenschaft des Repräsentationsmodus selbst – mithin *im Unterschied* zu Texten, Gemälden, Fotografien, etc. – zu erfassen vermag.

Hier rückt unmittelbar der Begriff der ‚Multimodalität‘ in den Blick, da dieser die dynamische Komplexität audiovisueller Bilder adressiert. Wird er doch im Kontext der Kognitionswissenschaften auf distinkte (menschliche) Ausdrucksmodi – wie Verbalsprache, Schriftsprache, Geste, piktorale Modi etc. – bezogen. In bildtheoretischer Sicht hingegen ist das audiovisuelle Bewegtbild als solches ein gänzlich eigener medialer Repräsentationsmodus, der eben dadurch ausgezeichnet ist, dass er alle möglichen Aspekte akustischer und visueller Wahrnehmungsdimensionen in einem gestalthaften Bewegungsbild vereinen kann.<sup>73</sup> Das audiovisuelle Bewegtbild ist also einerseits durch seine Multimodalität als spezifischer Ausdrucksmodus definiert. Andererseits bringt es zugleich eine genuine Weise multimodaler Wahrnehmungsfigurationen hervor, die als das Ganze eines Bewegungsbildes, eben als ‚multidimensionale Erfahrungsgestalt‘ zu begreifen ist, die sich nicht durch die Addition einzelner Sinnesmodalitäten erfassen lässt.

Die Frage nach den audiovisuellen Metaphern kann deshalb auch nicht auf ein ‚Mapping‘ zwischen semantischen Feldern etwa von sprachlichen Ausdrücken und ikonischen Repräsentationen heruntergebrochen werden; selbst wenn man das Mapping mehr oder weniger überzeugend auf kognitive Schemata rückbeziehen kann. Sie muss vielmehr immer die medienästhetischen Modellierungen

---

73 Zum Unterschied zwischen Bewegtbild und Bewegungsbild vgl. Fußnote 61 im ersten Teil dieses Buches.



rekonstruieren, in denen die Worte und Bilder, die Klänge und Bewegungsmuster sich Medienrezipienten präsentieren, um in der Performativität ihres Sich-selbst-Präsentierens, ihres Sich-zur-Wahrnehmung-Bringens als einheitliches Gestaltmuster ‚Sinn zu machen‘.

## 2.5 Grenzen und Möglichkeiten neurowissenschaftlicher Beschreibungsmodelle

Ich will mit meiner Kritik an den kognitionstheoretischen Modellen keineswegs das Konzept kognitiver Schemata generell zurückweisen. Mir geht es vielmehr um eine Differenzierung der unterschiedlichen Geltungsansprüche. Insofern gilt meine Kritik ausschließlich der Anwendung solcher Modelle als Taxonomien bedeutungsgenerierender Elemente oder affektiver Pattern, mit denen sich Inhalte audiovisueller Medien klassifizieren und analysieren lassen. Hier gilt der hinlänglich erörterte Einwand: Theorien, die auf einem individualpsychologischen Entwurf des psychischen Apparats gründen, können nur dann für die Analyse audiovisueller Bilder in Anspruch genommen werden, wenn man von der konkreten, historisch-kulturell situierten medialen Praxis absieht. Der Rezipient wird dann entsprechend dem je gewählten kognitiven Schematismus als bedeutungsgenerierender Automat modelliert – und alles, was sich nicht in die Schematismen des zugrunde gelegten Kognitionsmodells eintragen lässt, wird als historisch-kulturelles Wissen deklariert und additiv angefügt.<sup>74</sup>

Andererseits können neuropsychologische Konzepte für die theoretische Modellierung medialer Rezeptionsprozesse als temporale Muster der Verkörperung – d. i. gerade nicht als universelle Elemente der Bedeutungsgenerierung – instruierend sein. Dabei geht es nicht um die Validität solcher Theorien im Feld der Neurolinguistik oder der kognitiven Psychologie. Die kann hier überhaupt nicht angemessen diskutiert werden. Wohl aber lässt sich feststellen, dass entsprechende Modelle einem theoretischen Verständnis individueller Verkörperungsprozesse den Boden bereiten, die vorstellbar machen, wie das Gefühl für die sensuellen und affektiven Prozesse des eigenen Körpererlebens von Zuschauern zur Quelle metaphorischer Verknüpfungen werden kann.

Begriffe wie *image schema*, *action schema* und *motor schema* erlauben es nämlich durchaus, die physisch-sinnliche Verwicklung konkreter Rezipienten in die temporale Struktur audiovisueller Bewegtbilder als Prozess des Embodiment

---

<sup>74</sup> Vgl. die Kritik der neoformalistischen Filmtheorie David Bordwells im ersten Teil des Buches (S. 40–51).

*theoretisch zu konzipieren.* Mit ihnen lässt sich die dynamische Interaktion zwischen audiovisuellem Bild und Zuschauerkörper sinnfällig als Prozess permanenter Modellierung des körperlichen Zuschauerempfindens beschreiben.<sup>75</sup>

Man wird deshalb auf Konzepte wie etwa das der „embodied simulation“<sup>76</sup> (Gibbs) keineswegs verzichten wollen. Kann doch mit ihnen sinnfällig die Dynamik des Zuschauens als ein Prozess rekonstruiert werden, dessen Temporalität in der kompositorischen Struktur des audiovisuellen Bildes greifbar ist.

Die diskutierten Probleme der Anwendung kognitionstheoretischer Konzepte auf die Analyse audiovisueller Bilder laufen also keineswegs auf eine ‚nature vs. nurture‘-Kontroverse hinaus. Sie betreffen vielmehr die Art und Weise, in der die Relationen zwischen kognitiven Schemata und kulturellen Ausdrucksformen, Medientechniken, medialen Praktiken und symbolischen Formen konzipiert werden; sie betreffen die Metapher des Kognitionsautomaten und die Aporien einer universellen Tiefensemantik.

### **Ein neurowissenschaftliches Modell verkörpernder Wortverarbeitung**

Wie aber verhalten sich Medientechniken, Bedeutungssysteme und die poetischen Konzepte kultureller Ausdrucksformen zu den kognitiven Schemata menschlichen Denkens? Will man die damit angesprochenen historisch-kulturellen Differenzierungen tatsächlich auf die Ebene kognitiver Operationen beziehen, muss man die Relationen zwischen den Formen kultureller Sinnkonstruktion und dem physiologisch gegründeten Kognitionsprozess anders fassen als mit universellen Elementen der Bedeutungsgenerierung, gleichviel ob diese als generative Grammatik oder als Tiefensemantik gedacht wird. Sie lassen sich dann auch nicht in der Metapher des Kognitionsautomaten beschreiben.

Auch wenn wir uns in unserer Forschergruppe nicht auf experimentell-empirisch ausgerichtete Rezeptionsforschung stützen, möchte ich dennoch cursorisch auf einen neurolinguistischen Forschungskomplex eingehen, den sich Neurolinguistik und kognitive Anthropologie teilen und der eng mit der *CMT* verbunden

---

<sup>75</sup> Vgl. Hermann Kappelhoff/Jan-Hendrik Bakels: Das Zuschauergefühl. Möglichkeiten qualitativer Medienanalyse, in: *zfm – Zeitschrift für Medienwissenschaft* 5 (2), 2011, S. 78–95.

<sup>76</sup> Vgl. Gibbs: *Metaphor Interpretation as Embodied Simulation*; Gibbs: *Embodiment and Cognitive Science*; Raymond W. Gibbs: *Our Metaphorical Experiences of Film*, in: *Cinematic Metaphor in Perspective. Reflections on a Transdisciplinary Framework*, hrsg. v. Sarah Greifenstein, Dorothea Horst, Thomas Scherer, Christina Schmitt, Hermann Kappelhoff, Cornelia Müller, Berlin/Boston 2018.

ist.<sup>77</sup> Mehr noch, für diese Forschung wurde die These, dass in der Logik der Metaphernbildung kognitive Prozesse greifbar sind, die der Sprachbildung vorausgehen, zu einer ihrer Initialzündungen. Eröffnete sich doch damit die Möglichkeit, die Entstehung und Entwicklung menschlicher Sprachen zu erforschen, ohne von gegebenen sprachlichen Universalien auszugehen;<sup>78</sup> eine Möglichkeit, die wiederum von der strukturalistischen wie der poststrukturalistischen Kulturtheorie für lange Zeit als falsch gestellte Ursprungsfrage verabschiedet worden war.

Betrachtet man die Ergebnisse und theoretischen Implikationen dieser Forschung, wird deutlich, dass sich die Relationen zwischen kognitiven Prozessen, bedeutungsgenerierenden Systemen und medialen Techniken durchaus als nicht-repräsentationale Bezüge fassen lassen, die den Aporien des Dualismus von universeller generativer (Tiefen-)Struktur und bedeutungsgenerierender kognitiver

---

**77** Ich beziehe mich hier vor allem auf die Kritik sprachlicher Universalien (vgl. Stephen Levinson: *Presumptive Meanings. The Theory of Generalized Conversational Implicature*, Cambridge, Mass. 2000; Adele Eva Goldberg: *Constructions. A Construction Grammar Approach to Argument Structure*, Chicago 1995; Adele Eva Goldberg: *Constructions at Work. The Nature of Generalization in Language*, Oxford 2006) und die damit verbundene experimentelle Forschung von Adele Goldberg und Friedemann Pulvermüller zur Entstehung sprachlicher Ausdrucksformen. Im Hintergrund steht dabei die Forschung von Tomasello zur ‚kulturellen Entwicklung menschlichen Denkens‘ und zur menschlichen Empathiefähigkeit als Grundlage der Sprachentwicklung (vgl. Michael Tomasello: *Die kulturelle Entwicklung des menschlichen Denkens. Zur Evolution der Kognition*, Frankfurt/M. 2006; Michael Tomasello: *Warum wir kooperieren*, Berlin 2010). Vgl. zudem die Arbeiten von Eve Sweetser et al. zum Zusammenspiel gestisch-körperlicher und grammatikalisch strukturierter, sprachlich-deiktischer Bezugnahme auf multiple perspektivische Wirklichkeitsbezüge: Eve Sweetser: *What Does it Mean to Compare Language and Gesture? Modalities and Contrasts*, in: *Crosslinguistic Approaches to the Psychology of Language. Research in the Tradition of Dan Isaac Slobin*, hrsg. v. Jiansheng Guo et al., New York 2009, S. 357–366; Eve Sweetser: *Introduction: Viewpoint in Language and Gesture from the Ground Down*, in: *Viewpoint in Language. A Multimodal Perspective*, hrsg. v. Barbara Dancygier, Eve Sweetser, Cambridge 2012, S. 1–24; Eve Sweetser/Lilian Ferrari: *Subjectivity and Upwards Projection in Mental Space Structure*, in: *Viewpoint in Language. A Multimodal Perspective*, hrsg. v. Barbara Dancygier, Eve Sweetser, Cambridge 2012, S. 47–68; Eve Sweetser: *Creativity Across Modalities in Viewpoint Construction*, in: *Language and the Creative Mind.*, hrsg. v. Michael Borkent et al., Stanford 2013, S. 239–254; Eve Sweetser/Kashmiri Stec: *Maintaining Multiple Viewpoints in Gaze*, in: *Viewpoint and the Fabric of Meaning. Form and Use of Viewpoint Tools across Languages and Modalities*, hrsg. v. Barbara Dancygier, Wei-lun Lu, Arie Verhagen, Berlin 2016, S. 237–257.

**78** Vgl. Nicholas Evans/Stephen Levinson: *The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science*, in: *Behavioral and Brain Sciences* 32 (5), 2009, S. 429–448; siehe auch Goldbergs Kommentar zu diesem Artikel: Adele Eva Goldberg: *Essentialism Gives Way to Motivation*, in: *Behavioral and Brain Sciences* 32 (5), 2009, S. 455–456. Vgl. auch Adele Eva Goldberg: ‚Origins of human communication‘. By Michael Tomasello. Reviewed by Adele E. Goldberg, in: *Language* 85 (4), 2009, S. 953–955.

Operation entgehen. Mit ihnen wird ein Modell der ‚kulturellen Entwicklung des Denkens‘<sup>79</sup> absehbar, das in der Logik metaphorischer Konzeptbildungen ein grundlegendes Paradigma der Dynamik dieser Entwicklung findet.

Ich möchte diesen Gedanken mit Blick auf eine experimentelle Versuchsreihe aus der Neurolinguistik ausführen; die Studie zielt in ihrer theoretischen Begründung und in ihrer Fragestellung auf die Prozesse der Verkörperung, die von der *CMT* allen metaphorischen Konzeptbildungen unterlegt werden. Es handelt sich um Experimente, die Prozesse der Wortverarbeitung untersuchen. Sie stützen sich in ihren Beobachtungen weitgehend auf die Bildgebungsverfahren der Magnetresonanztomografie (MRT) – versuchen also, Hirnaktivitäten bestimmter Areale zu identifizieren, zu messen und zu qualifizieren. Die Frage war, ob sich Relationen zwischen Verben und den motorischen Aktivitäten feststellen lassen, die von diesen Verben benannt werden. D. h., gibt es Indizien, dass eine benannte Aktivität auf der neuronalen Ebene mit dem entsprechenden körperlichen Vollzug verbunden ist? Gewählt wurde ein Wort („Greifen“), das auf einen klar umrissenen Bewegungsablauf verweist, dessen Zuordnung innerhalb der entwickelten Topografie MRT-gestützter Hirnforschung unstrittig ist. Getestet werden sollte die rezeptive Aktivität beim Hören des Wortes ‚Greifen‘ (Wortverarbeitung) im Verhältnis zur entsprechenden motorischen Aktivität. Im Ergebnis konnte festgestellt werden, dass die neuronalen Aktivierungsmuster, die der perzeptiven Verarbeitung der gehörten Lautfolge des gesprochenen Wortes ‚Greifen‘ zuzuordnen sind, mit einer Aktivierung jener neuronalen Muster einhergingen, die von der MRT-gestützten Hirnforschung als Aktionsschemata der motorischen Aktivität des Greifens identifiziert worden sind.

Tatsächlich hat sich in der Studie die hypothetisch angenommene Korrelation bestätigt. Auf der Ebene der MRT zeigt sich eine zeitliche Interdependenz zwischen der neuronalen Aktivität, die der Verarbeitung des Wortes zuzuordnen ist, und dem neuronalen Aktionsschema der motorischen Aktivität ‚Greifen‘. Die Verarbeitung der gehörten Lautfolge des Wortes ‚Greifen‘ lässt sich auf dieser Ebene also als ein neuronales Aktivierungsmuster (perzeptives Schema) beschreiben, das mit dem Aktionsschema ein zeitliches Intervall (Perzept-Aktion) bildet.

Die gleiche Verknüpfung lässt sich zeigen, wenn das Wort ‚Begreifen‘ zu hören gegeben wird.<sup>80</sup> D. h., auf der Beschreibungsebene der MRT ist die Verarbeitung

---

<sup>79</sup> Tomasello: Die kulturelle Entwicklung des menschlichen Denkens.

<sup>80</sup> Véronique Boulenger/Yury Shtyrov/Friedemann Pulvermüller: When Do You Grasp the Idea? MEG Evidence for Instantaneous Idiom Understanding, in: *Neuroimage* 59 (4), 2012, S. 3502–3513; Friedemann Pulvermüller: Meaning and the Brain: The Neurosemantics of Referential, Interactive, and Combinatorial Knowledge, in: *Journal of Neurolinguistics* 25 (5), 2012, S. 423–459; Frie-

des wörtlich zu verstehenden Verbs nicht von der metaphorischen Übertragung zu unterscheiden: Beides gründet auf dem Intervall zwischen neuronalem Perzeptions- und Aktionsschema.

Keineswegs will ich mit diesem Exkurs die experimentelle neurolinguistische Forschung als empirische Begründung eigener Thesen in Anspruch nehmen. Doch lassen die Ergebnisse dieser und vergleichbarer Studien<sup>81</sup> einige theoretische Schlüsse über die Verknüpfung bedeutungsbildender Prozesse mit sensorischen Schemata ihrer Verkörperung zu, die ich etwas genauer skizzieren möchte. Mein Ziel ist es dabei, die augenfälligen theoretischen Annahmen, die den Experimenten zugrundeliegen resp. sich aus ihrer Auswertung ergeben, als ein Beschreibungsmodell von Verkörperungsprozessen zu rekonstruieren, auf das sich die Analyse metaphorischer Sinnkonstruktionen ohne essentialistische Setzungen beziehen kann. Mir geht es also nicht um die Beweiskraft neurolinguistischer Experimente (eine Frage, die nur in der entsprechenden Fachdisziplin zu klären ist), sondern um das darin implizite Beschreibungsmodell.

Die Ausgangshypothese der genannten Experimente ist, dass symbolische Entitäten, seien dies Worte, Bilder oder Zeichen, auf der neuronalen Ebene genauso wie Dinge als Wahrnehmungen verarbeitet werden. Die neuronale Verarbeitung des gehörten Wortes ‚Greifen‘ betrifft dann zunächst nichts anderes als die akustische Wahrnehmungssensation der Lautfolge. Zum Baustein eines bedeutungsgenerierenden Konzeptes kann das Verarbeitungsschema dieser Lautfolge nur dann werden, wenn es mit der neuronalen Aktivierung des Aktionsschemas ‚Greifen‘ verknüpft wird.

Die Vermutung liegt nahe, dass die Verknüpfung – oder das Intervall – selbst gleichsam als semantischer Index fungiert, der die Perzeption der bloßen Lautfolge als symbolische Form, eben als Wort markiert. Das setzt wiederum voraus, dass das perzeptive Schema der Lautfolge des Wortes ‚Greifen‘ in regelmäßig wiederholten Verknüpfungen mit dem Aktionsschema ‚Greifen‘ als neuronales Intervall perpetuiert worden ist. Es setzt weiterhin voraus, dass das Aktionsschema von der motorischen Aktivität getrennt wird, die mit ihm eigentlich initiiert ist: die physisch realisierte Bewegung des Greifens.

Um die Lautfolge ‚Greifen‘ zu etwas werden zu lassen, das als Wort verarbeitet wird, muss also einerseits das Intervall von Perzeptions- und Aktionsschema

---

demann Pulvermüller/Clare Cook/Olaf Hauk: Inflection in Action: Semantic Motor System Activation to Noun- and Verb-Containing Phrases is Modulated by the Presence of Overt Grammatical Markers, in: *Neuroimage* 60 (2), 2012, S. 1367–1379.

<sup>81</sup> Vgl. die Beiträge in Diane Pecher/Rolf A. Zwaan (Hrsg.): *Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking*, Cambridge 2005.

durch wiederholte Verknüpfung perpetuiert, andererseits die motorische Umsetzung des Aktionsschemas blockiert werden. Auf der Beobachtungsebene der MRT lässt sich folglich ein zeitliches Intervall zwischen Wortverarbeitung und Aktivierung motorischer Schemata ausmachen, das mit einer Modifikation des Aktionschemas verbunden ist. Es bleibt virtueller Bewegungsimpuls.

Man könnte sagen, dass die Verknüpfung mit dem perzeptiven Verarbeitungsschema des Wortes ‚Greifen‘ das Aktionsschema in die Schwebe eines Bewegungsimpulses versetzt, der auf der physischen Aktion insistiert, ohne dass diese motorisch umgesetzt wird. Ich möchte mit Blick auf den medientheoretischen Kontext meiner Argumente diesen insistierenden Impuls einer physischen Aktion, die virtuell bleibt, als *Affekt* bezeichnen.<sup>82</sup> Dann ließe sich tatsächlich in einem strengen Sinne von Netzen interagierender Perzeptions-, Affektions- und Aktionsschemata sprechen, deren Intervalle sich zu sensomotorischen Schemata kognitiver Konzepte verbinden. Nur lassen sich solche Vernetzungen sinnvoll weder als Binnenstruktur psychischer Apparate beschreiben, noch aus einer solchen herleiten.

Wenn das einzige Kriterium, mit der die perzeptive Verarbeitung der Lautfolge als Wortverarbeitung qualifiziert werden kann, in dem Umstand besteht, dass die Lautfolge ‚Greifen‘ ein Aktionsschema aktiviert, welches selbst – bezogen auf seine motorische Umsetzung – modifiziert wird, dann ist davon auszugehen, dass die perpetuierende Wiederholung der Verknüpfung sich auf das vollständige Intervall bezieht. Die Ausgangsform muss als Intervall von perzeptiver Verarbeitung, aktiviertem Aktionsschema und motorischer Ausführung gedacht werden; während die symbolische Form eine Modifikation darstellt, in der das letzte Glied, in die Schwebe einer virtuellen Aktion versetzt, *nicht mehr getan*, sondern *nurmehr gedacht* wird.

### **Kooperation und Kognition**

Eine solche Genese aber ist nur als Interaktion zwischen einer Pluralität von Akteuren vorstellbar; die Lautfolge (‚Greifen‘), die zum Aktionsimpuls (‚Zugreifen‘) wird, ist sinnfällig nur als koordinierende Expression in der Interaktion voneinander geschiedener, aber aufeinander bezogener Körper vorstellbar. Dann nämlich ist die Lautfolge ‚Greifen‘ nichts anderes als eine intentionale (Laut-)Geste (ein Befehl, ein Ruf, eine Aufforderung), die einen anderen (einen hörenden, sehen-

---

<sup>82</sup> Das ist präzise die Definition, die Deleuze in seiner Auseinandersetzung mit Bergson herausarbeitet; vgl. Deleuze: Das Bewegungs-Bild, S. 97 und S. 123 ff.

den, rezipierenden) Körper zu dirigieren sucht. Sie ist Expression eines Körpers, der eine unbestimmte Vielheit eigensinniger Körper koordinierend auf eine intentional gleichgerichtete Aktion ausrichtet. Und umgekehrt: Sie ist impulssetzende Wahrnehmung, die eine motorische Aktivität initiiert, mit der sich ein rezeptiver Körper einfügt in eine koordinierte Aktion. Das Wort ‚Greifen‘ ginge also auf die intentionale Expression eines Körpers zurück, der in einem anderen Körper die motorische Aktion Greifen initiiert.

Insofern wäre auf der neuronalen Ebene gar nicht zwischen dem Wort ‚Greifen‘ und der entsprechenden Aktion zu unterscheiden; beides wäre durch das gleiche Intervall von initiiender Expression, (re-)perzeptivem Schema, Aktionsschema und motorischer Umsetzung bestimmt. Von dorther betrachtet wäre lediglich zwischen dem perpetuierten sensomotorischen Intervall einer koordinierten (Inter-) Aktion einerseits zu unterscheiden, das im letzten Glied des Intervalls der motorischen Ausführung den einzelnen Körper einfügt in eine koordinierte Aktion einer Pluralität von agierenden Körpern; und einem sensomotorischen Konzept andererseits, in dem das gleiche Intervall von initiiender Sensation (Perzept) und initiiertes Aktion, in seiner motorischen Ausführung blockiert oder gehemmt, zum reinen Bewegungsimpuls (Affekt) wird; die Aktion selbst wird nicht (mehr) vollzogen, sondern *als sensomotorisches Szenario nur noch gedacht, d. h., sie wird als Handlungsmöglichkeit reflektiert*. Freilich ist das, was da nur noch gedacht und nicht mehr getan wird, immer schon komplex.

Kognitive Konzepte lassen sich in dieser Perspektive zwar auf grundlegende sensomotorische Schemata zurückführen; beziehen sich aber auf sensomotorische Schemata, die immer schon alles das umfassten, was notwendig mit der koordinierten Aktion gegeben sein muss; sie umfassten ein raum-zeitliches Aktionsfeld, das durch die *perspektivische Verortung einer Vielheit* aufeinander sich abstimmender Akteure und deren koordinierte intentionale Ausrichtung auf *eine Objektwelt* strukturiert ist (und nicht durch universelle Raum-Zeit-Schemata).

Man ist versucht, mit Blick auf solche sensomotorischen Schemata von multidimensionalen Erfahrungsgestalten und deren metaphorischer Verknüpfung zu sprechen; das erklärte dann auch, warum *auf dieser Beobachtungsebene* die wörtliche Bedeutung von ‚Greifen‘ kategorial nicht von der metaphorischen Konzeptbildung im Wort ‚Begreifen‘ zu unterscheiden ist. Der Begriff würde die Erweiterung des Intervalls von Aktions- und Perzeptionsschema um eine reflexive Rückkopplung erfordern.

Aber die sensomotorischen Schemata, von denen hier die Rede ist, sind im Abstraktionsgrad auf der gleichen Ebene angesiedelt wie die *image schemas*; sie geben lediglich eine skelettierende Beschreibung der Erfahrungskonstellation, die der medientechnisch extrem vermittelten Beobachtungsebene experimenteller Neurolinguistik entlehnt ist. Im Unterschied freilich zu den *image schemas*

beziehen sie sich immer schon auf ein multiperspektivisch strukturiertes Netz von Perzeptions-, Affektions- und Aktionsschemata, das dem Aktionsfeld untereinander sich koordiniert verhaltender Körper zugehört. Was wiederum bedeutet, dass sich eigensinnige Körper in ihrer Vielheit aufeinander abstimmend beziehen und sich als ein ‚Wir‘ gegenüber der Objektwelt anderer menschlicher und nicht-menschlicher Organismen und Körper verhalten können.

Tatsächlich sieht die hier diskutierte Forschung den Grund für die Entwicklung der menschlichen Sprachen in einem Merkmal der Spezies, das Michael Tomasello mit seiner Forschung zur kognitiven Anthropologie prominent gemacht hat: Die menschliche Fähigkeit zu kooperativem Handeln.<sup>83</sup> Tomasello hat mit seiner experimentellen Beobachtung des Verhaltens nicht-menschlicher Primaten und Kleinkinder eine Perspektive auf die Entwicklungsgeschichte menschlichen Denkens eröffnet, die den Gegensatz von biologischem Essentialismus und kulturtheoretischem Konstruktivismus zu überwinden sucht. Keine universelle biologische Mitgift der Spezies, sondern eine Verhaltensmodifikation, die sich in einer empathischen Bezugnahme auf andere menschliche Lebewesen bekundet, wird zum anthropologischen Differenzkriterium.

Wenn man schon bei Kleinkindern empathiegeleitetes kooperierendes Verhalten beobachten kann, das über das Kalkül der intelligenten Beutegemeinschaft hinausgeht, darf man davon ausgehen, dass die Entwicklung menschlichen Denkens insgesamt von dieser Verhaltensmodifikation geprägt ist. Die Fähigkeit zur Empathie schafft die Voraussetzung für ein kooperatives Handeln, das in sich die Notwendigkeit und Möglichkeit der Sprachbildung birgt.<sup>84</sup> Letztlich gründet sie sich auf dem Umstand, dass das ‚Wir‘ kooperierender Interaktionen eine dynamische Ausdehnung der Selbstwahrnehmung impliziert, ein sich in den Relationen von ‚Ich‘, ‚Wir‘ und ‚die Anderen‘ immer neu refigurierendes Selbst.<sup>85</sup>

Sensomotorische Schemata, die als kognitive Konzepte fungieren, sind immer schon eingelassen in die kooperierende Interaktion von Körpern eines dynamisch sich verändernden ‚Wir‘. Deshalb ist das Netz der Perzeptions-, Affektions- und Aktionsschemata nicht als Binnenstruktur eines psychischen Apparates in der Metapher des Kognitionsautomaten zu beschreiben. Es müsste vielmehr als Intervall-Diagramm des relationalen Gefüges einer dynamischen Konfiguration von Körpern vorgestellt werden, die koordiniert agieren; sie sind funktional auf

---

<sup>83</sup> Vgl. Tomasello: Warum wir kooperieren.

<sup>84</sup> Vgl. Michael Tomasello: Eine Naturgeschichte des menschlichen Denkens, Berlin 2014.

<sup>85</sup> Entwicklungspsychologisch wurden von Winnicott bis Stern die Formen solcher dynamischen Selbstfigurationen, die auf empathischen Beziehungsmustern aufrufen, in der Interaktionen der Mutter/Kind-Beziehung erforscht.



die permanente Refiguration der raum-zeitlichen Dimensionen des Aktionsfeldes und der perspektivischen Verortung der Akteurspositionen bezogen.

Menschliches Wahrnehmen und Denken ist immer situiert; es verhält sich in einer dynamisch sich verändernden Konstellation des Mit-Seins von Menschen, Lebewesen und Dingen, die in abgestuften Graden vertraut, gekannt, verstanden oder fremd sind und deshalb in höchst unterschiedlicher Intensität affizieren und Projektionen triggern. Von den grundlegendsten kognitiven Verknüpfungsschemata bis zu den komplexesten „grammatikalischen Perspektivierungen“ (Eve Sweetser) ist das menschliche Denken und Wahrnehmen in die kooperative Interaktion aufeinander empathisch abzustimmender Körper eingelassen.

### **Kognitive Schemata und kulturelle Differenzierung**

Dann aber erscheint Bedeutungsproduktion selbst – Verstehen und Reflektieren in der konkreten Interaktion – unmittelbar mit einem Prozess der Ausdifferenzierung je vorgängiger kognitiver Schemata einherzugehen. Sie hat immer schon Anteil an einer Entwicklungsgeschichte des Denkens, die auf der Ebene des Verhaltens und Lernens als kulturelle Entwicklung sich vollzieht. Medien, Sprachen und semiotische Systeme konstituieren jene Tiefendimension kultureller Sinnressourcen, in der das Denken sich aus den je konkreten Interaktionen herauslösen und als kooperierende Aktion reflektieren, d. i., sich auf sich selbst beziehen und selbst beschreiben kann. Die permanente Differenzierung kognitiver Schemata wäre somit schlicht auf der Ebene der Medienpraxis, des Gebrauchs von Techniken der kooperativen Interaktion und der Reflexion zu verorten. Sie wäre selbst Teil kultureller Sinnproduktion. Medientechniken, semiotische Systeme und Poetiken kultureller Ausdrucksformen markieren in dieser Perspektive kulturgeschichtliche Plateaus der Ausdifferenzierung des menschlichen Denkens.

Man mag einwenden, dass ich hier höchst abstrakte Hypothesen neurolinguistischer Forschung als ein konkretes Modell kultureller Differenzierungsprozesse umformuliere. Tatsächlich erhebe ich nicht den Anspruch, die von mir thematisierte Forschung in ihrem fachwissenschaftlichen Wert einzuordnen und zu qualifizieren. Mir sei dieser Ausflug in die spekulative Sprachursprungstheorie trotzdem erlaubt; verfolge ich damit doch keinen anderen Zweck, als ein der neurolinguistischen Forschung implizites Beschreibungsmodell des Wechselverhältnisses zwischen den kognitiven Schemata und dem Gebrauch symbolischer Systeme (hier die Wortsprache) zu extrapolieren, das auf andere mediale Techniken der Sinnproduktion übertragen werden kann. Letztlich war das Ziel, eine andere Metapher als die des Kognitionsautomaten zu rekonstruieren, mit der die

dynamische Funktion metaphorischer Konzepte für die kulturelle Entwicklung menschlichen Denkens in den Blick gerückt wird.

Das sollte zumindest insofern erreicht worden sein, als dass mit dem rekonstruierten Beschreibungsmodell metaphorische Sinnkonstruktionen filmischer Bewegungsbilder als Prozesse verkörpernder Simulation (*embodied simulation*) sinnfällig zu beschreiben und zu erklären sind – ohne auf universelle kognitive Strukturen oder Systeme zurückzugreifen.

Hier nun schließt sich der Kreis. Ist doch aus meiner Sicht die Metapher eines solipsistischen Kognitionsautomaten der Grund dafür, dass die *CMT* als angewandte Kulturanalyse sich in den Aporien einer physiologisch begründeten generativen Tiefensemantik verstrickt. Damit wird jene Entdeckung verspielt, die am Anfang der *CMT* stand: die Möglichkeit, in der Konstruktionslogik metaphorischer Konzepte die Sprachentstehung selbst als einen Prozess kultureller Differenzierung erforschen zu können.

Indessen macht gerade die Digitalisierung aller menschlichen Interaktion greifbar, dass es sehr viel sinnvoller ist, die extensiven Dynamiken medialer Modellierungsprozesse menschlichen Denkens in den Blick zu nehmen, als nach universellen Gesetzmäßigkeiten zu suchen, auf die alle Modulationen zurückzuführen sind. Die „metaphors we live by“ mögen sehr viel eher ihren funktionalen Sinn in der Dynamik der kulturellen Entwicklung menschlichen Denkens, der Adaption an die rasend schnell sich verändernden medialen Bedingungen menschlichen Zusammenlebens finden als in einer generativen Semantik.

Das ist der Grund, warum wir in unserer Forschung am Konzept der ‚multidimensionalen Gestalt‘ festgehalten haben, auch wenn es nur schwer mit der neuropsychologischen Terminologie zu vermitteln ist, die die kognitive Medien- und Metaphernforschung gegenwärtig dominiert. Lassen sich doch multidimensionale Erfahrungsgestalten unmöglich auf eine Taxonomie feststehender kognitiver Schemata bringen. Sie sind immer schon eingelassen in die multiperspektivische Anordnung kooperativer (Inter-)Aktionsfelder.