

Hartmut Winkler

# Ähnlichkeit, Schemabildung, Medien.

## 1. Der Skandal der Ähnlichkeit

Ähnlichkeit – und da Sie die Ähnlichkeit zum Thema Ihres Kolloquiums machen, sage ich Ihnen hiermit nichts Neues – ist eine überaus interessante, gleichzeitig aber schwierige Kategorie. Dies gilt zunächst für den Alltag, in dem die Wahrnehmung von Ähnlichkeiten allgegenwärtig ist und alle Sphären durchdringt; noch mehr aber für Philosophie und Kulturwissenschaften, die die Ähnlichkeit – hier ist Dorothee Kimmich zuzustimmen – lange Zeit für nicht theoriefähig hielten.<sup>1</sup> Und wenn sich dies gegenwärtig zu ändern beginnt, insofern Sie, oder ich und auch andere die Ähnlichkeit aus der Sicht des jeweiligen Faches noch einmal neu in den Blick nehmen, dann zeigt dies einen Umbruch an.

Ähnlichkeit als Kategorie ist verwirrend und irritierend. Was ähnlich ist, ist nicht identisch und eben auch nicht völlig different; eine Art Zwischenzone, die eigentümlich vage erscheint.



Und häufig kann man nicht einmal genau sagen, in welcher Hinsicht ein Ding dem anderen ähnlich oder eben unähnlich ist. Die Feststellung von Ähnlichkeit – hört man – sei eine Sache der Intuition; und ganz sicher gilt dies in der Sphäre der Bilder, weshalb es notorisch schwierig ist, Algorithmen für die Bildersuche zu schreiben. Ähnlichkeit setzt Entferntes in Relation; und Assoziationen, die Pfaden der Ähnlichkeit folgen, können Abgründe überspringen.

Im Alltag haftet der Ähnlichkeit nicht selten etwas Unheimliches an. Zwillinge irritieren unsere Vorstellung von Individualität, indem sie unseren Blick spalten und Verwechslungen provozieren. Aus dem Spiegel blickt uns morgens manchmal ein Fremder an. Der frühe Film hat, in Auseinandersetzung mit dem Realismus der filmischen Abbildung, die Figur des Doppelgängers gefeiert. Elvis-Impersonators schaffen es, Elvis ähnlicher als Elvis zu sein. Und selbst die industrielle Serienproduktion kann etwas Unheimliches haben; wenn z. B. Einfamilienhäuser einander allzu ähnlich sind, geht uns das auf die Nerven.



2

<sup>1</sup> Kimmich, Dorothee: *Ins Ungefähre. Ähnlichkeit und Moderne*. Paderborn: Brill/Konstanz University Press 2017; Bhatti, Anil; dies.: *Ähnlichkeit. Ein kulturtheoretisches Paradigma*. Paderborn: Brill/Konstanz UP 2015.

<sup>2</sup> <https://www.spiegel.de/wirtschaft/tuerkei-in-der-geisterstadt-burj-al-babas-stehen-500-schloesser-leer-a-1250518.html>

Vor allem aber tut sich die Theorie mit dem Thema überaus schwer. Verkürzt kann man sagen, dass die Ähnlichkeit dem Grundgestus aller Theorie widerspricht, die Dinge auseinanderzulegen, zu analysieren und das heißt zu trennen.<sup>3</sup> Theorie, und speziell die letzten 50 Jahre poststrukturalistischer Theoriebildung haben in radikaler Weise die Differenz favorisiert. Zunächst in Frontstellung gegen die ‚Identität‘, die, wie gesagt, den zweiten Pol der Ähnlichkeit bildet, letztlich aber auch gegen die Ähnlichkeit selbst.



(Allerdings könnte man alternativ die These vertreten, dass ‚Differance‘ und Wiederholung gleichzeitig den Blick auf die Ähnlichkeit lenken...)

Die Karriere des Digitalen spiegelt am deutlichsten wider, dass sich Wissenschaft immer noch an Descartes‘ Ideal eines ‚clara et distincta‘ orientiert. Ähnlichkeit kann dem nicht gerecht werden. Ähnlichkeit ist und bleibt eine ‚weiche‘, um nicht zu sagen ‚schmutzige‘ Kategorie.

Warum ist Ähnlichkeit dann aber dennoch so interessant? Ich möchte im Folgenden in verschiedene Richtungen argumentieren. Zum einen ist mir wichtig, dass Ähnlichkeit allgegenwärtig ist; zumindest im Feld der Medien, sodass mein Fach, die Medienwissenschaft, eigentlich gar keine andere Wahl hat, als Ähnlichkeit zu einem ihrer Gegenstände zu machen. Damit ist auch schon die Perspektive markiert, aus der ich zu Ihrem Projekt beitragen will: Mein Blick ist ausschließlich auf die Medien gerichtet.

Zum zweiten möchte ich in einem einzelnen Beispielfeld zeigen, dass Ähnlichkeit möglicherweise eben doch theoriefähig ist. Theoriefähig zumindest in dem Sinne, dass Theorie ohne die Kategorie der Ähnlichkeit nicht auskommen kann. Und theoriefähig möglicherweise auch, wenn gezeigt werden kann, dass die Ähnlichkeit – implizit – sehr etablierte Theoriefelder immer schon dominiert. Als Beispielfeld habe ich die Schemabildung gewählt; was dies bedeutet und was ihr Stellenwert ist, wird, so hoffe ich, im Folgenden deutlich werden.

## 2. Wahrnehmung

Bereits die menschliche Wahrnehmung verfährt nach Kriterien der Ähnlichkeit. In der Zeit um 1900 haben die Gestalttheorie, aber auch Wahrnehmungspsychologen wie Stern oder Ebbinghaus betont, dass wir große Schwierigkeiten hätten, Dinge überhaupt wahrzunehmen, sie gegen ihren Hintergrund freizustellen, wenn wir sie nicht *wiedererkennen* würden. So beschreibt Ebbinghaus die Situation eines Säuglings, der in seinem Wagen liegt:

„Ein ganz junges Kind blicke von einer bestimmten Stelle aus in ein bestimmtes Zimmer. Es empfängt davon einen wenig gegliederten, diffusen Eindruck. Nun werde es von der Mutter in seinem Wagen in ein benachbartes Zimmer geschoben; in der Hauptsache tritt dann ein anderer Gesamteindruck an die Stelle des ersten. Aber die Mutter und der Wagen sind doch dieselben geblieben. Die von ihnen ausgehenden optischen Reize finden also die ihnen mögliche materielle und seelische Wirkung bereits etwas vorbereitet [...]; den übrigen, abgeänderten Teilreizen fehlt diese zwiefache Begünstigung. [...] Der vom Anblick (der Mutter) herrührende Eindruck kommt damit einerseits immer leichter zustande, andererseits reißt er sich immer mehr los von den

<sup>3</sup> Ein Durchgang durch verschiedene philosophische Ansätze findet sich ebenfalls bei Kimmich (K., D.: *Ins Ungefähre...*, a. a. O., S. 9-66).

verschiedenen diffusen Hintergründen, in denen er ursprünglich aufging: die Anschauung der Mutter wird ein immer selbständigeres Glied des jeweiligen Gesamteindrucks.“<sup>4</sup>

Für die Frage nach der Ähnlichkeit ist hier Folgendes wichtig: Zunächst, dass Objekte der Wahrnehmung nicht einfach gegeben sind, sondern erst langsam, Schritt für Schritt gegen ihren Hintergrund freigestellt und als Objekte konstituiert werden. Und damit eng verbunden: Dass sich die Eindrücke erst in der *Wiederholung* stabilisieren.

Wiederholung aber bedeutet Ähnlichkeit. Eindrücke, die einander ähnlich sind, gewinnen Kontur und schichten sich – in Spannung zu den wechselnden Kontexten – zu festen Wahrnehmungskonzepten auf. Dies erlaubt in der Folge ein *Wiedererkennen*; die Wahrnehmung ist nicht mehr ausschließlich mit Neuem konfrontiert, sondern kann in Anschlag bringen, worüber sie aus Erfahrung bereits verfügt. Dies schließt ein, dass in den Prozess der Wahrnehmung das *Gedächtnis* immer schon involviert ist; mit der Folge, dass die jeweils aktuelle Wahrnehmung keineswegs nur eine aktuelle Wahrnehmung ist.

Auf Basis dieser frühen Wahrnehmungspsychologien nun hat man ausgebaute *Schematheorien* entwickelt. Ausgearbeitet wurden diese vor allem von der Kognitiven Psychologie; und in der Sache durchaus nicht unproblematisch.<sup>5</sup> Das Konzept des Schemas selbst aber ist ein großer Gewinn; zum einen weil es, wie ich zeigen werde, ausgehend von der Wahrnehmung auch in andren Feldern verwendbar ist; und zum zweiten, weil es – ähnlich eben wie die Ähnlichkeit – eine gewisse Unschärfe immer schon zugesteht. Schemata, das sagt schon der Begriff, fallen mit dem Schematisierten nicht einfach zusammen. Sie sind abstrakter als das Schematisierte; und gleichzeitig würde man damit rechnen, dass sie zwar stabil, gleichzeitig aber immer abhängig von Entwicklung, immer im Umbau sind.

Wahrnehmung also *sucht* Ähnlichkeit. Sie beobachtet Ähnlichkeit und extrahiert sie aus dem amorphen Material, das die Sinne bieten. Und dies beginnt außerhalb der menschlichen Sphäre schon in der Natur, weil viele Tiere nicht nur über angeborene, sondern auch über erworbene Schemata verfügen. Der Prozess der Schemabildung selbst ist angelegt und verläuft vollständig automatisiert; und er geht über die Objekterkennung weit hinaus, insofern auch zeitliche Vorgänge in Prozess-Schemata abgelegt werden; und schließlich gibt es einen Übergang zu nahezu allen anderen Typen von Regularitäten. Schemabildung, könnte man sagen, ist der Basismechanismus, auf den letztlich auch *Abstraktion* und *Form* zurückgeführt werden können. Gehen wir von der Wahrnehmung nun zu den Medien, und zunächst zur Sprache über.

### 3. Medium Sprache

In der Sprache – ihrer inneren Struktur und ihrer Funktionsweise – spielt die Ähnlichkeit eine herausragende Rolle, und dies auf den unterschiedlichsten Ebenen; so. z. B. wenn man Kontiguität und Similarität unterscheidet, oder wenn de Saussure der syntagmatischen Achse die ‚assoziativen‘ Reihen gegenüberstellt, die einer Logik der Ähnlichkeit folgen;<sup>6</sup> oder aber in den verschiedenen Dimensionen sprachlicher Mimesis, die ebenfalls immer Ähnlichkeit implizieren.

Aus all diesen Möglichkeiten möchte ich – exemplarisch – nur eine einzelne herausgreifen und näher verfolgen, und dies ist der Mechanismus der *Subsumtion*, der für das Funktionieren

<sup>4</sup> Ebbinghaus, Hermann: Grundzüge der Psychologie. Bd. 2, Leipzig 1913, S. 15; zit. nach: Kittler, Friedrich A.: Aufschreibesysteme 1800 - 1900. München: Fink 1985, S. 326f.

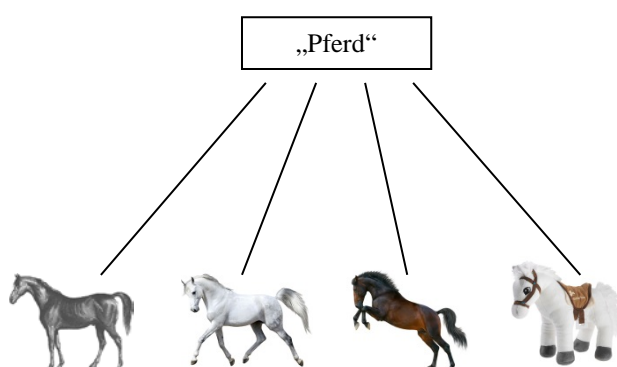
<sup>5</sup> Winkler, Hartmut: Schemabildung – Eine Maschine zur Umarbeitung von Inhalt in Form. In: Conradi, Tobias u. a. (Hg.): Schemabildung und Praktiken. München: Fink 2012, S. 15-35.

<sup>6</sup> De Saussure, Ferdinand: Grundfragen der allgemeinen Sprachwissenschaft. Berlin 1967, S. 153ff. (EV., frz.: 1916).

der Sprache, die Semantik und die Begriffsbildung von absolut grundlegender Bedeutung ist. Wie Begriffe entstehen, ist durchaus strittig. Die linguistische Theorie würde üblicherweise sagen, dass Worte/Begriffe durch Unterscheidung zustande kommen und dass der Wortschatz sich nach den Regeln von Kontrast und Differenz gliedert.<sup>7</sup> Ich werde zeigen, dass Ähnlichkeit eine mindestens ebenso große Rolle spielt; zum einen weil Unterscheidung, Differenz und Kontrast nicht einfach das Andere der Ähnlichkeit sind, sondern auf Ähnlichkeit und sogar auf ‚Identität‘ zurückbezogen bleiben.

Und zum zweiten ganz direkt, insofern Begriffe immer und grundsätzlich subsumieren. Der Begriff ‚Pferd‘, um ein Beispiel zu nehmen, verweist nicht auf ein einzelnes Exemplar, sondern auf eine Kategorie oder Gattung, die die unterschiedlichsten Exemplare umfasst; und dies gilt analog, wenn wir Namen ausklammern, für alle Begriffe und Worte.

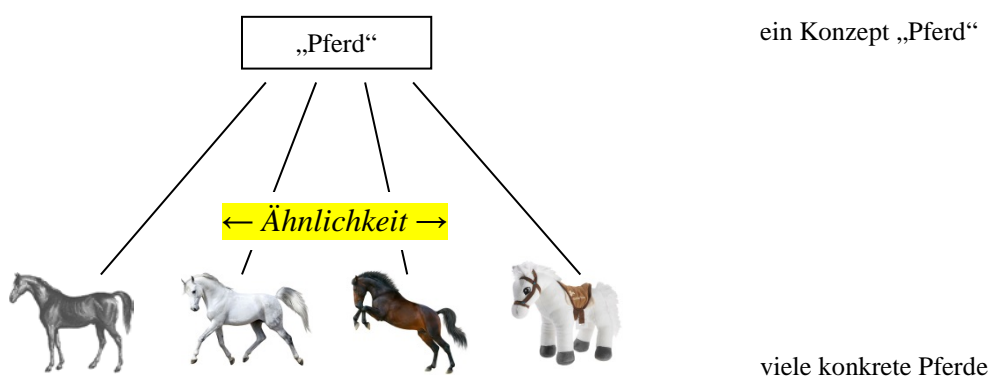
Sprache, Literatur:



Hierin liegt einer der entscheidenden Unterschiede zwischen Sprache und Welt, den Zeichen und der Sphäre der Referenten. Und zwar qualitativ wie auch im Hinblick auf Quantitäten: Einer unendlich großen Zahl von Exemplaren steht ein relativ begrenztes Set von Begriffen gegenüber; Voraussetzung dafür, dass möglicherweise nicht die Welt, sehr wohl aber die Sprache in unsere kleinen Köpfe passt.

Entscheidend nun ist, dass die Exemplare einander *ähneln* müssen, damit sie unter einen gemeinsamen Begriff fallen.

Sprache, Literatur:



Die Bildung des Konzepts ‚Pferd‘ *verdankt sich der Ähnlichkeit*. Eine Sprachgemeinschaft beobachtet, dass bestimmte Tiere einander besonders ähnlich sind; und diese Beobachtung

<sup>7</sup> Lyons, John: Semantik. Bd. 1, München 1980, S. 281-330 (EV., am.: 1977).

hält sie fest, indem sie diese zu einer Gruppe zusammenfasst und dann mit einem Gruppen-Etikett, der sprachlichen Bezeichnung ‚Pferd‘, versieht.

Geht man ins Detail, werden die Dinge komplizierter. Bei der Subsumtion nämlich werden nur ganz bestimmte Beobachtungen, bestimmte Ähnlichkeiten berücksichtigt; diejenigen Merkmale eben, die nach Ansicht der Sprachgemeinschaft ein Pferd zu einem Pferd machen. In anderer Hinsicht können die Exemplare ausgesprochen *unähnlich* sein, sich also stark unterscheiden. Im Fall ‚Pferd‘ etwa ist die Farbe irrelevant; das Merkmal wird für die Begriffsbildung unterdrückt.

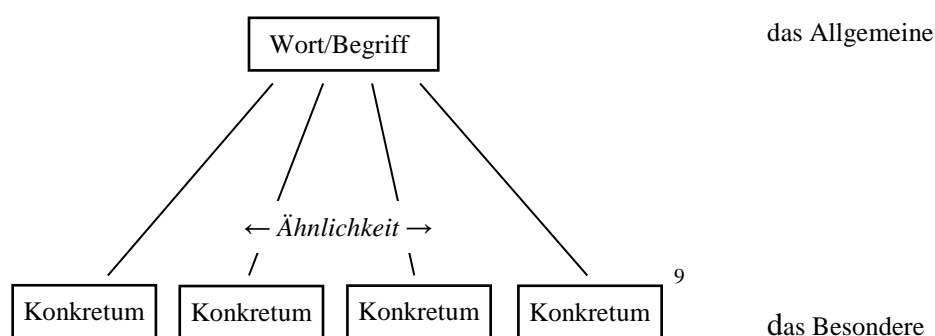
Die Beobachtung anderer Ähnlichkeiten wird entsprechend zu völlig anderen Gruppenbildungen führen: Kaninchen und Steine etwa würden sich drastisch unterscheiden, d. h. ganz überwiegend *unähnlich* sein; und dennoch können sie das Merkmal ‚grau‘ teilen. Die Bezeichnung, der Begriff ‚grau‘ hält diese eine, sehr partikuläre Ähnlichkeit fest.



Ein Pferd nun kann ein Pferd, und gleichzeitig grau sein. Dies macht deutlich, dass die Kategorien sich überlagern, sodass man in der Semantik von einem komplexen Geflecht sich überlagernder Ähnlichkeiten und Kontrastachsen ausgehen muss. Weiter wird man festhalten müssen, dass Semantik nicht darauf angewiesen ist, dass es sich bei den Ähnlichkeiten immer um visuelle handelt; und schließlich gibt es viele abstrakte Begriffe ohne direkten Weltbezug, die sich nicht oder nur sehr vermittelt auf eine vor-geordnete Welt stützen. Nicht umsonst hat Walter Benjamin die Sprache deshalb ein System ‚*unsinnlicher Ähnlichkeiten*‘ genannt.<sup>8</sup>

Halten wir fest, dass die einzelnen Begriffe subsumieren; und dass dies immer Ähnlichkeit impliziert.

Sprache, Literatur:



<sup>8</sup> Benjamin, Walter: Lehre vom Ähnlichen. In: Ges. Schriften., Bd. II/1, Frankfurt a. M. 1980, S. 204-210, hier: S. 207 (EV.: 1933).

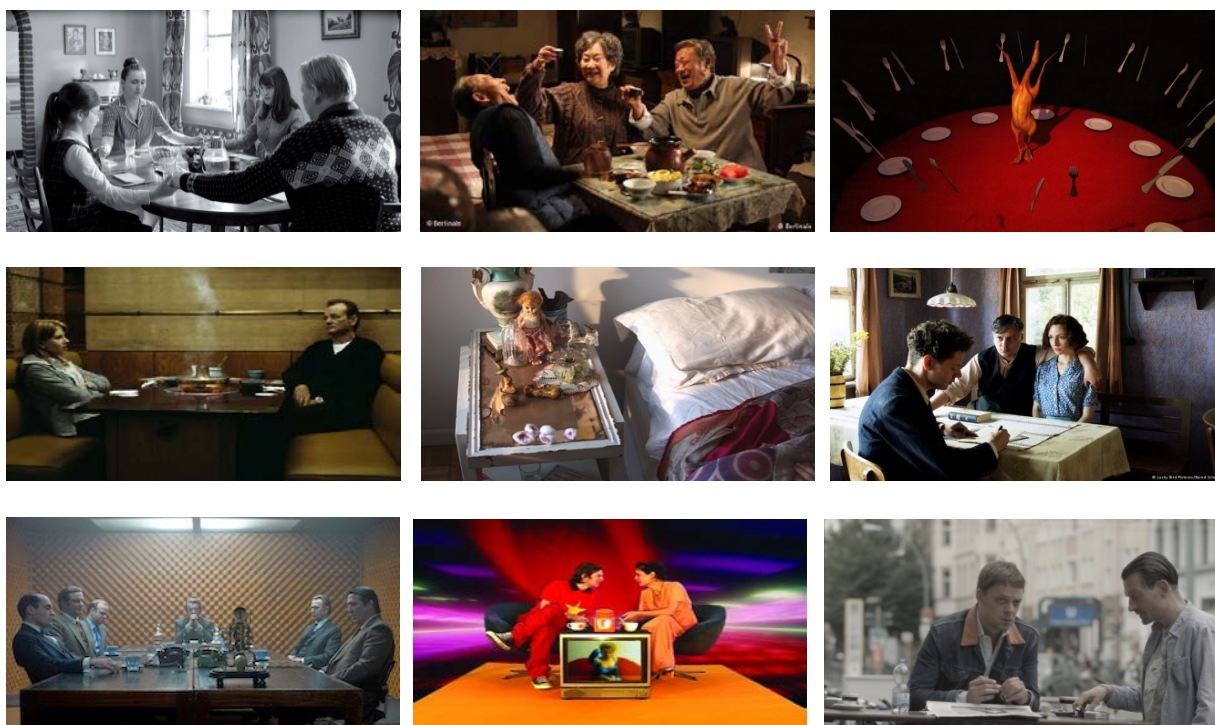
<sup>9</sup> Wenn ich hier von Konkreta spreche, dann in abgeschliffenem Sinn; als Fachterminus meint ‚Konkretum‘ bereits einen Begriff, einen Begriff allerdings, der – im Gegensatz zu Abstrakta – etwas Gegenständliches, konkret sinnlich Erfahrbares bezeichnet. (Vgl.: Wikipedia, dt.: Konkretum; <http://de.wikipedia.org/wiki/Konkretum>).



#### 4. Bildmedien: Fotografie und Film

Als zweiten Medienkomplex möchte ich die technischen Bilder, Fotografie und Film, ansprechen, gerade weil diese auf den ersten Blick völlig anders als Sprache und Schrift funktionieren. Die augenfälligste Dimension von Ähnlichkeit ist die Ikonizität, die Tatsache, dass Bilder, anders als Worte, dem Abgebildeten ‚ähneln‘; nicht darum aber soll es mir gehen, vielmehr möchte ich auch hier – beispielhaft isoliert – nur das Problem der Subsumtion diskutieren.

Fotografie und Film nämlich haben ihre Pointe darin, dass sie ohne Begriffe auskommen und den Mechanismus der Subsumierung vermeiden. Wo die Sprache das Wort ‚Tisch‘ verwendet, also ein allgemeines Konzept, das zunächst nur eine relativ abstrakte Vorstellung evoziert und noch sehr viele, unterschiedliche Tische bezeichnen könnte, präsentieren Fotografie und Film das Abbild jeweils eines *einzelnen* Tisches. Die Bilder von Fotografie und Film liefern *Konkrete*.



Sprache und Literatur müssen zum Mittel der Beschreibung greifen, um Begriffe zu konkretisieren; und das heißt, dem Begriff weitere an die Seite stellen. Die Bilder dagegen liefern die Deskription immer schon mit; über die Farbe und Beschaffenheit des fotografierten Tisches ist immer schon disponiert; es ist vollständig unmöglich, dass auch nur eine dieser Bestimmungen im Vagen bleibt.

Gleichzeitig *schwächen* die Bilder die Kontraste und Grenzziehungen, die in der Sprache garantiert schienen, und eröffnen das Spiel der Ähnlichkeiten neu. Fotografie und Film unterlaufen das identifikatorische Denken; Wolken können wie Pferde aussehen, ein Pferd wie ein Schatten, ein Schatten kann einem Fleck ähneln, und zwar mehr als jedem anderen Schatten... Mediengeschichtlich bedeuten Fotografie und Film eine Revision; und zwar gerade an der Front jener Distinktionen und Distinktionskraft, die in der Sprache immer schon gewährleistet schienen. Dies ist die eine Seite der technischen Bilder.

Doch sie haben noch eine andere, unvermutete Seite, und hier wird es für die Ähnlichkeit endgültig interessant. Das Gesagte nämlich heißt nicht, dass jedes technische Bild tatsächlich ein Unikat wäre, oder dass jedes fotografierte Objekt nur für sich selbst stünde. Denn selbstver-

ständig können wir Fotos oder Filme nur dann verstehen, wenn wir den Tisch – bei aller Konkretheit – als einen Tisch erkennen oder wiedererkennen.

In der Konsequenz heißt dies, dass das identifikatorische Denken auch hier eine Rolle spielt; und damit wieder die Ähnlichkeit. Wir können den Tisch nur dann als Tisch identifizieren, wenn wir über ein Konzept, ein Schema, das Stereotyp eines Tisches bereits verfügen. Die Kognitionstheorie sagt uns, dass wir ein ganzes Set, ein *System von Schemata und Stereotypen* in uns tragen, das uns überhaupt erst erlaubt, der Welt gegenüberzutreten und mit der Welt umzugehen. Den Schatz eines Vorwissens, in dem Medienkompetenz und Realerfahrungen sich mischen.

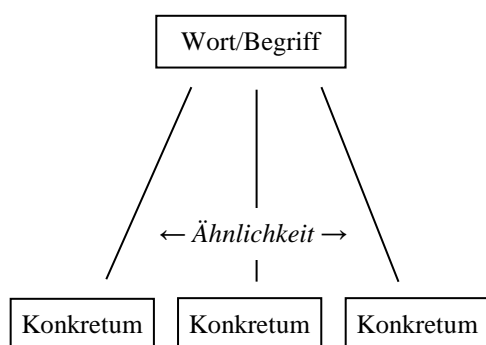
Anzuerkennen, wie zentral die Rolle dieser Schemata ist, fällt z. B. der Filmwissenschaft keineswegs leicht. Meist wird schlicht vorausgesetzt, dass es Schemata gibt, und häufig werden diese pejorativ bewertet; und ebenso gilt für selbstverständlich, dass es sich bei Filmbildern um ‚gegenständliche‘ Bilder handelt. In der bildenden Kunst ist dies anders. Der Einbruch der Abstraktion in die Kunst und übrigens auch der experimentelle Film haben gezeigt, wie groß die Verwirrung ist, wenn es sich eben nicht mehr um gegenständliche Bilder handelt, wenn die Bilder das Objekterkennen und unseren Satz etablierter Schemata gezielt unterlaufen. Das System der Schemata in die Analyse der Bildmedien einzubeziehen ist von zentraler Bedeutung; eben weil ohne sie das Medium nicht funktioniert und nicht erklärt werden kann, wie und warum wir Bilder verstehen.

Hinter den differenten Oberflächen der Bilder also walten die Schemata und damit – die Ähnlichkeit. Das Set der Schemata und Stereotypen bildet eine zweite Ebene, die – wenig sichtbar – hinter den Bildern liegt. Und diese Schemata stehen, insofern Schemata immer *abstrakt* sind, *in einer systematischen Spannung zur Konkretheit der Bilder*.

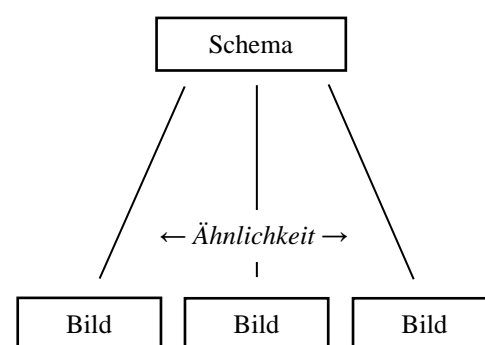
## 5. Medienvergleich

Wenn wir nun auf das zurückgehen, was wir zur Sprache gesagt haben, ergibt sich eine verblüffende Parallele: Die Schemata/Stereotypen nämlich nehmen eine ähnliche Position ein, wie sie im Fall der Sprache die Begriffe haben:

Sprache, Literatur:



Bildmedien, Fotografie, Film:

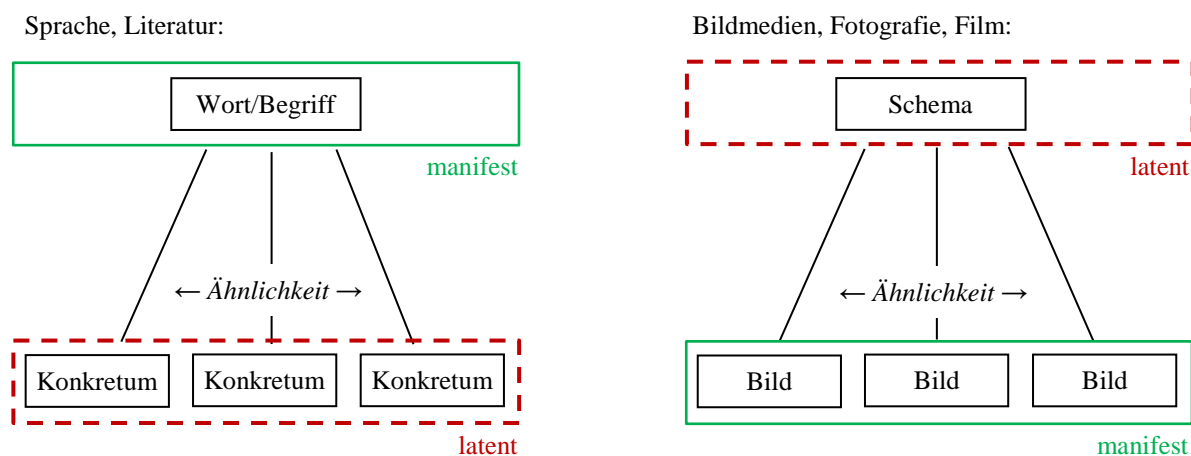


Und wie im Fall der Worte/Begriffe ist es wieder die Ähnlichkeit, die den Sprung vom jeweils Einzelnen zum Schema erlaubt. Sofort aber werden auch entscheidende Differenzen deutlich; und zwar vor allem in Bezug auf die Frage, was die beiden Ebenen trennt und welchen Stellenwert sie jeweils haben.

Im Fall von Sprache und Literatur sind es zunächst die Texte, die materiell auf dem Tisch liegen und den Gegenstand der Untersuchung bilden. Und da Texte aus Worten/Begriffen gemacht sind, ist es die obere Ebene in meiner Skizze, die materiell manifestiert und beobacht-

bar ist. Die Begriffe mit konkreten Vorstellungen zu füllen dagegen bleibt dem Leser überlassen; hier ist seine Imagination gefordert; wenn im Text ein Pferd vorkommt, muss er sich vorstellen, um was für ein Pferd es sich handelt und welche Farbe es hat. Diese Konkretisierung bleibt im Kopf des Rezipienten und damit latent.

Im Fall von Fotografie und Film ist es genau umgekehrt: Da die materiell manifestierten Bilder, wie gesagt, die *Konkreta* liefern, sind sie auf der unteren Ebene der Skizze zu finden. Und entsprechend sind es hier gerade die Schemata, die latent bleiben, weil der Rezipient sie – als Medienkompetenz oder Weltwissen – mitbringen oder bereitstellen muss.



Wie, in aller Welt, aber kann es zu einer solchen Differenz kommen? Vor allem für die Ähnlichkeit: Wieso kann die Ähnlichkeit einmal die Ebene des Latenten, also die vorgestellten Konkreta, und einmal die materiellen, manifesten Bilder betreffen?

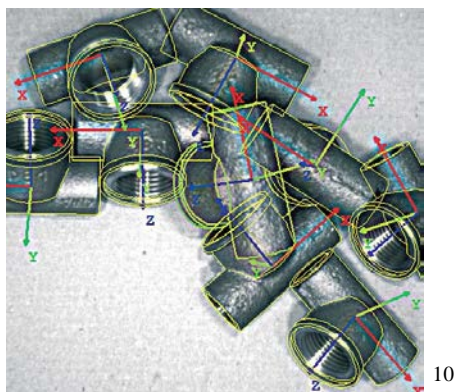
Meine Antwort wäre, dass exakt in solchen Differenzen der Grund dafür liegt, dass es überhaupt verschiedene Medien gibt. Möglicherweise wichtiger als die übliche Weise, Sprache von Bildern, begrifflich von visuell, und ikonische von arbiträren Zeichen zu unterscheiden, wäre damit, wie die Medien mit Ähnlichkeit umgehen, ihre Schemabildung, ihren Typus von Ähnlichkeit organisieren. Auf diesen Punkt werde ich in meinem Schlussteil noch einmal zurückkommen.

## 7. Computer, maschinelle Objekterkennung

Ein Schritt aber fehlt noch, bevor dies passieren kann. Denn möglicherweise erscheint Ihnen meine Überlegung zum Medienvergleich als schematisch/abstrakt; oder schlimmer: als eine Art Glasperlenspiel, das man, wenn man nicht gerade Medienwissenschaftler ist, ebenso gut auch lassen kann. Ich möchte deshalb zumindest kurz noch ein drittes Medium ansprechen, weil dieses nahezu alles, was ich bisher gesagt habe, auf einen Punkt zusammenzieht und zeigt, dass es sich um eine sehr konkrete und praktisch relevante Frage handelt. Und dieses Medium ist der Computer, bzw. – damit wir im Feld der Bilder bleiben – eine seiner Anwendungen, die maschinelle Objekterkennung.

Objekterkennung wird z.B. in der industriellen Produktion eingesetzt, etwa wenn ein Roboterarm nach einem Maschinenteil greifen soll. Zur Steuerung verwendet man Videokameras; und dann braucht man ein Programm, das hilft, das Teil, egal wo es im Haufen liegt, zu finden.

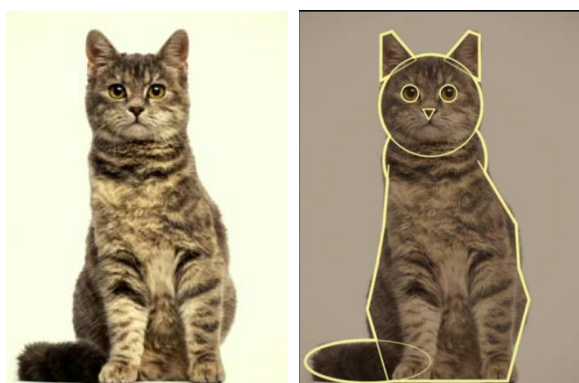




10

Schwieriger wird es, wenn es um eine größere Anzahl *unterschiedlicher* Dinge geht; etwa wenn man prüfen will, ob ein Bild einen Hund oder eine Katze zeigt. Diese Aufgabe, die einem dreijährigen Kind spielend leicht fällt, ist für einen Computer ein ernstes Problem. Zunächst müssen aus dem kontinuierlichen Bild, das für sich genommen ja keine Objekte kennt, ‚Objekte‘ überhaupt erstmal herausgestanzt werden; und dann geht es darum, die Objekte zu identifizieren; und hierfür muss der Rechner unterscheiden können; er muss über Kategorien verfügen, denen er die Objekte jeweils zuordnen kann.

Wie also geht man eine solche Aufgabe an? In einem ersten Schritt hat man versucht, für alle Objekte Listen typischer Merkmale zusammenzustellen. Wenn etwas einen runden Kopf und spitze Ohren hat, ist es möglicherweise eine Katze.



11

Dieser Ansatz allerdings ist schnell gescheitert, weil sich die konkreten Abbildungen stark unterscheiden und kaum ein Bild alle Merkmale erfüllt, die man für unabdingbar halten würde.



12

<sup>10</sup> <https://www.computer-automation.de/feldebene/bildverrbeitung/artikel/72779/2/>

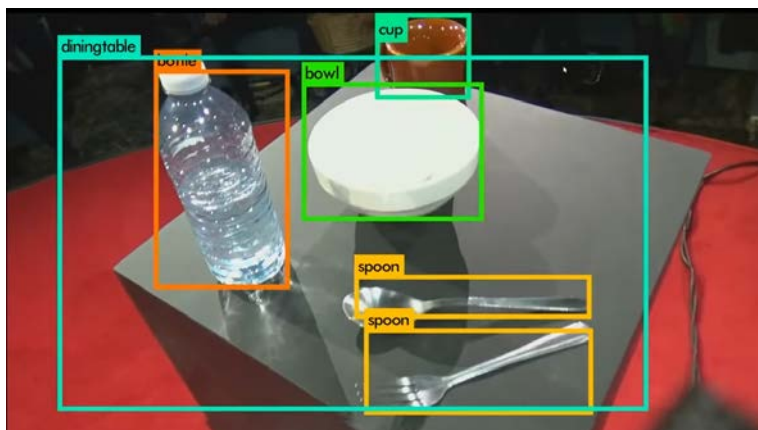
<sup>11</sup> Video: Fei Fei Li: How we teach computers to understand pictures;  
<https://www.youtube.com/watch?v=40riCqvRoMs> , Min: 05:00ff.

<sup>12</sup> Ebd.

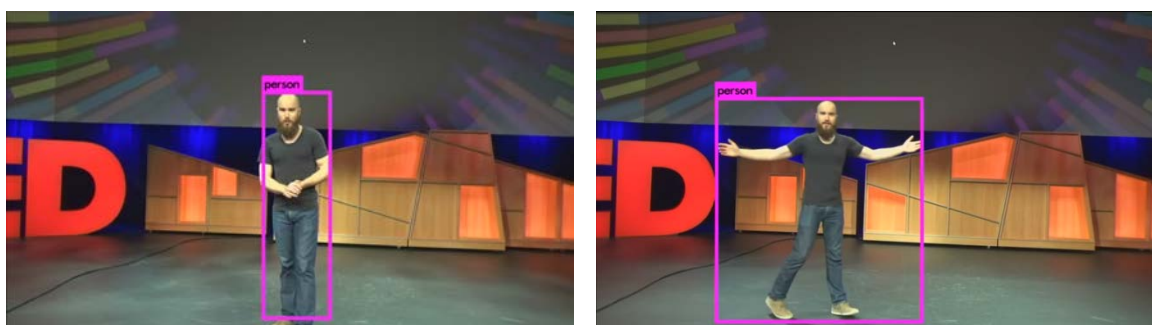
Deshalb ging man einen anderen Weg und setzt inzwischen das sogenannte ‚Deep Learning‘ und ‚künstliche neuronale Netzwerke‘ (Convolutional Neural Networks) ein.<sup>13</sup> Diese Programme haben die Besonderheit, dass sie auf die Definition expliziter Merkmale völlig verzichten; stattdessen werden die Programme ‚trainiert‘, indem man sie mit großen Mengen bereits vorhandener Bilder füttert. Hierfür sind sehr viele Lernzyklen nötig und der Prozess zerfällt – grob – in drei Phasen: Zu Beginn müssen Menschen dem Netzwerk sagen, um welche Objekte es sich jeweils handelt, was man ‚Labeling‘ nennt; in einer zweiten Phase greift man nur noch dann korrigierend ein, wenn das Netzwerk falsche Ergebnisse liefert; so dass in der dritten das Programm schließlich selbstständig läuft und – ohne weiteren Eingriff des Menschen – Objekte in Bildern erkennt; oder nur noch so viele Fehler macht, wie die Macher für vertretbar halten.

Die Programme sind auf eine extrem große Datenbasis angewiesen. Das bisher ambitionierteste Projekt ‚Imagenet‘ hat 2009 eine Datenbank aufgebaut, die Millionen von Bildern enthält und diese nach immerhin 22.000 Kategorien ordnet. Diese Datenbasis wurde mit Hilfe von 50.000 Crowd-Sourcing-Mitarbeitern – also händisch, von Menschen – erstellt. Und da sie öffentlich zugänglich ist, greifen viele der Projekte zur automatischen Bilderkennung auf diese Daten zurück, um ihre Programme zu trainieren.

Bilddatenbanken und neuronale Netze haben die Bilderkennung in einem Maße voran gebracht, das noch vor wenigen Jahren kaum jemand für möglich gehalten hätte. So können die Programme, legt man ihnen Fotos vor, unterschiedliche Objekte identifizieren.



Sie erkennen auch solche Objekte, die sich in Bewegung befinden, oder ihre Größe verändern.



14

Und noch wesentlich komplizierter ist der dritte Fall: die Steuerung autonomer Fahrzeuge durch den komplexen Verkehr einer Stadt. Hier muss der Rechner eine Vielzahl von Objekten identifizieren; und zwar auch dann, wenn diese in Bewegung, perspektivisch verzerrt oder

<sup>13</sup> Wikipedia: [https://de.wikipedia.org/wiki/Convolutional\\_Neural\\_Network](https://de.wikipedia.org/wiki/Convolutional_Neural_Network)

<sup>14</sup> Video: How computers learn to recognize objects instantly | Joseph Redmon, <https://www.youtube.com/watch?v=Cgxsv1riJhI> , Min: 03:10ff.

teilweise abgedeckt sind; und da Autofahren eine zeitkritische Angelegenheit ist und Kleinkinder ziemlich plötzlich auf die Straße laufen, muss das Ganze in Realzeit geschehen.



15

All das ist ziemlich eindrucksvoll, auch wenn viele Probleme nach wie vor ungelöst sind und es noch eine Weile dauern wird, bis wir vollautomatische Fahrzeuge auf den Straßen sehen. Warum aber gehört die maschinelle Objekterkennung in den hier von mir verfolgten Zusammenhang?

Zum einen ist augenfällig, dass es auch hier um Ähnlichkeit geht. Maschinen sollen in die Lage versetzt werden, Ähnlichkeit zu erkennen, eine Fähigkeit, die man bis dahin ausschließlich Menschen zugetraut hätte. Und wenn es gelingt, diese Fähigkeit tatsächlich in Algorithmen zu packen, muss man folgern, dass Ähnlichkeit vielleicht eben doch theoriefähig ist.

Zum zweiten imitieren die Programme die menschliche Wahrnehmung, und zwar vor allem, insofern sie Wahrnehmung an Wiedererkennen binden. Auch hier gibt es die beiden Ebenen, die ich oben gezeigt habe: Es gibt die jeweils aktuellen Wahrnehmungen und es gibt ‚darüber‘ die Schemata (auch wenn diese hier Kategorien heißen). Und wie oben für die technischen Bilder beschrieben geht es darum, die Konkretion der Bilder hinter sich zu lassen und in heterogenem Material Schemata aufzufinden.

Ein wesentlicher Unterschied allerdings besteht darin, dass den Künstlichen Neuronalen Netzwerken die Kategorien vorgegeben werden, während sie im Fall der Wahrnehmung – das war der Punkt bei Kittler und Ebbinghaus – im Prozess wiederholter Wahrnehmung aus dem Material emergieren.<sup>16</sup> Die Wahrnehmung bildet ihre Kategorien im Prozess der Wahrnehmung selbst, das können Künstliche Neuronale Netzwerke nicht.

## 8. Folgerungen

Ich gehe damit zu meiner Schlussüberlegung über. Denn wie nun – so wird man fragen müssen – gehören die verschiedenen Bausteine, die ich diskutiert habe, zusammen? Gibt es einen Übergang zwischen den Mechanismen der Wahrnehmung und dem Funktionieren der Medien? Der Rolle der Ähnlichkeit in der Gestalterkennung und der Begriffsbildung der Sprache oder den Schemata im Bilderdiskurs? Ich denke, dass man einen solchen Übergang tatsächlich zeigen kann. Und ich denke, dass mein Computerbeispiel dies schon ein Stück weit veranschaulicht hat.

<sup>15</sup> [https://www.youtube.com/watch?v=\\_kxX09i4fds](https://www.youtube.com/watch?v=_kxX09i4fds)

<sup>16</sup> Dieser Unterschied ist sehr weitreichend. Im Kern geht es darum, dass sich das Material im Prozess wiederholter Wahrnehmung quasi automatisch ordnet; die Schemata und Kategorien bilden sich in der Wiederholung selbsttätig heraus. In der Welt der Rechner setzt man hierfür z. B. Clustering-Algorithmen ein...

Setzen wir noch einmal bei der Wahrnehmung an. Wahrnehmung, hatte ich gesagt, sucht gezielt nach Ähnlichkeit. Wenn alles Wahrnehmen Wiedererkennen ist, an die Feststellung von Ähnlichkeit, an Wiederholung und Gedächtnis gebunden, und wenn schon unser Wahrnehmungsapparat über eine Schemabildung verfügt, dann liegt es nahe, hier eine Brücke zu den Schemata des Bilderdiskurses zu schlagen.

Ganz offenbar sind es ähnliche Mechanismen, die Mutter Natur zunächst unserem Wahrnehmungsapparat eingeschrieben hat, und die die menschliche Geschichte dann als ‚Medien‘ im Raum der Kultur installiert. Oder – plausibler –: möglicherweise *imitieren* und verlängern die Medien, was in den Mechanismen unserer Wahrnehmung vorgeformt ist.

Wenn dies der Fall ist, dann allerdings – dies muss sofort ergänzt werden – auf signifikante Weise verschoben: Medien operieren im intersubjektiven Raum, im Raum des Kollektivs, und das heißt: *zwischen* den Menschen, weshalb sie häufig vorschnell auf ‚Kommunikation‘ reduziert werden. Und sie operieren im Außenraum, manifest/materiell, insofern sie immer auch Technik in Anspruch nehmen. Will man Kurzschlüsse vermeiden, also wird es nötig sein, den Zusammenhang etwas auszubuchstabieren.

Wenn Wahrnehmung Ähnlichkeit sucht, hat dies ökonomische Gründe. Es geht darum, die Komplexität der Welt soweit zu reduzieren, dass sie ihren überwältigenden Charakter verliert und zumindest einigermaßen handhabbar wird; und wenn der Wahrnehmungsapparat Schemata ausbildet, die ihm ein Wiedererkennen von Objekten erlauben, dann verhindert er, dass er überschwemmt wird, wie er zweifellos überschwemmt würde, würde er nur mit absolut Neuem konfrontiert werden.

*Exakt die gleiche Funktion* – dies ist meine Behauptung – *haben die Medien*. Meine Überlegung zur Subsumtion hat gezeigt, dass auch hier eine Reduzierung das Ziel ist. Es gelingt, eine sehr große Zahl von Einzelexemplaren unter einen einzelnen Begriff, ein einzelnes Schema zu bringen; mit der Folge, dass anschließend Begriff oder Schema an deren Stelle treten, sodass – ungleich ökonomischer – nur noch mit Begriff oder Schema gearbeitet werden muss. Auch hier geht es um den Aufbau eines Vorwissens, das es uns ermöglicht, mit dem jeweils Neuen umzugehen.

Ähnlichkeit ist die Bedingung und das Gesetz, das diese Reduzierung erlaubt. Ihre Feststellung ist immer perspektivisch, reduktionistisch, selektiv und prekär, weshalb Begriffe und Schemata grundsätzlich ‚lügen‘; Medien, die ohne diese Reduktion operieren, aber gibt es nicht. Medien sind in ihrem Funktionieren an die Schemabildung gebunden.

In den Mittelpunkt des Medialen – und man wird dies hervorheben müssen, weil es auch in meinem Fach leider keineswegs Konsens oder Standard ist – rücken damit Schema und *Zeichen*. Nicht Kommunikation oder Technik, nicht die einzelnen Medien, die in ihrer Verschiedenheit immer wieder die Aufmerksamkeit auf sich ziehen, nicht die Regeln der Nachrichtentechnik, die ‚Information‘ oder ‚das Digitale‘, sondern allein die gemeinsame/übergreifende Eigenschaft, mit Zeichen zu operieren, macht die Medien aus.

Wieso aber ‚Zeichen‘, wenn bislang doch nur von Schemabildung die Rede war? Meine letzte Überlegung nun soll auch diesen Schritt noch plausibel machen. Die Schemabildung bereits, das war an der Subsumtion deutlich geworden, ist ein Prozess der *Typisierung* und der *Verhärtung*. Die Feststellung von Ähnlichkeit favorisiert nicht nur bestimmte Aspekte und schließt das Unähnliche aus, sondern schafft auch eine neue Entität – das Schema – eben, das als ein Quasi-Ding in die Welt kommt und ein Eigenleben beginnt.

*Und wenn die Typisierung und Verhärtung weit vorangeschritten ist*, das ist mein Vorschlag, *würde man von einem ‚Zeichen‘ sprechen*. Zeichen halten fest und sistieren – eben in quasi-dinghafter Form –, was einmal Urteile und Hypothesen über Ähnlichkeiten waren.

Das erlaubt es, diese Quasi-Dinge – die Zeichen – in den Diskurs zurückzuspielen und im Diskurs mit ihnen zu operieren. Als Platzhalter, als konventionalisiert-verhärtet-vorgefasste Urteile über Ähnlichkeit fungieren sie wie ein Fertigteil: Alle Urteile über Ähnlichkeit, die im Zeichen schon getroffen sind, müssen live im Diskurs nicht mehr getroffen werden. Und gleichzeitig eben, das ist die Dialektik, werden die Hypothesen über Ähnlichkeit, die das Zeichen enthält, in jeder neuen Verwendungssituation neu erprobt.

Mein Vorschlag also ist, den Begriff des Schemas ernster als den des Zeichens zu nehmen und Zeichen als eine Form von Schemata anzusehen, die besonders verhärtet sind.

Die Bestimmung bewährt sich, insofern sie nicht nur den Brückenschlag zwischen Wahrnehmung, technischen Bildern und den Zeichen/Begriffen der Sprache erlaubt, sondern sie erlaubt es auch, den *Unterschied* dieser Medien zu fassen: denn möglicherweise gibt es die technischen Bilder, Fotografie und Film, nur, weil sie eben *weniger verhärtet* sind als die sprachlichen Zeichen. Indem sie Konkreta präsentieren, zeigen sie Schemabildung *in actu*; und sie machen den Prozess sichtbar, den man Begriffen und Zeichen nicht mehr ansehen kann.

Medien, das wäre meine Schlussthese, organisieren Ähnlichkeit. Oder genauer:

1. *Medien sind gesellschaftliche Maschinen, die Komplexität reduzieren, indem sie Ähnlichkeit beobachten, schematisieren / typisieren und dem Diskurs typisiert – als Zeichenset – wieder zur Verfügung stellen.*
2. *Medien bannen das Unheimliche, das der Ähnlichkeit anhaftet, durch die Schaffung von Zeichen als Quasi-Objekten, durch Konventionalisierung und Reifizierung.*
3. *Und gleichzeitig stellt jeder neue Text, jedes neue Bild die Frage neu und setzt das Spiel zwischen ähnlich und unähnlich, Identität und Differenz auch gegen die etablierten Zeichen immer wieder in Gang.*