

Zsuzsanna Kondor

### Representations and Cognitive Evolution. Towards an Anthropology of Pictorial Representation

2011

<https://doi.org/10.25969/mediarep/16582>

Veröffentlichungsversion / published version  
Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Kondor, Zsuzsanna: Representations and Cognitive Evolution. Towards an Anthropology of Pictorial Representation. In: *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft*. Themenheft zu Heft 14, Jg. 7 (2011), Nr. 2, S. 48–71. DOI: <https://doi.org/10.25969/mediarep/16582>.

#### Erstmalig hier erschienen / Initial publication here:

<http://www.gib.uni-tuebingen.de/image/ausgaben-3?function=fnArticle&showArticle=229>

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under a Deposit License (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual, and limited right for using this document. This document is solely intended for your personal, non-commercial use. All copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute, or otherwise use the document in public.

By using this particular document, you accept the conditions of use stated above.

Zsuzsanna Kondor

## **Representations and Cognitive Evolution. Towards an Anthropology of Pictorial Representation**

### **Abstract**

Im vorliegenden Beitrag werde ich für die Möglichkeit einer evolutionären Anthropologie des Bildes argumentieren. Ich werde mich dabei einerseits auf den Begriff der kognitiven Evolution Merlin Donalds stützen und auf der anderen Seite auf den Enaktivismus. Die beiden Ansätze können insofern als divergent betrachtet werden, als Donalds Theorie auf der Evolution repräsentationaler Fähigkeiten basiert, während Enaktivismus (sowohl in konservativer als auch in radikaler Perspektive) die Notwendigkeit mentaler Repräsentationen mindestens in bestimmten Fällen anzweifelt. Doch diese Divergenz führt, wie ich zeigen möchte, nicht in einen Widerspruch, vielmehr erhellen beide die Bedeutung des extrakraniellen Kontexts sowie der Darstellungstechniken, die Ideen, Ängste usw. manifestieren. Zunächst werde ich den speziellen Charakter eines Bildes, das als ein Bild von etwas konstruiert ist, erläutern; dann werde ich die notwendigen kognitiven Fähigkeiten der Wahrnehmung und Darstellung besprechen und schließlich die Bedeutung bildlicher Darstellung für die kognitive Evolution nachweisen.

In this paper I will argue for the possibility of a kind of evolutionary anthropology of images. I will rely on the notion of cognitive evolution as Merlin Donald construed it, on the one hand, and on the enactive account of vision, on the other hand. The two approaches can be seen as divergent, since the Donaldian theory is based on the evolution of representational capabilities,

while enactivism (both conservative and radical wings) challenges the necessity of mental representation, at least in certain cases. But, as I suggest, this divergence does not lead to a contradiction, rather both illuminate the importance of the extra-cranial context, and the importance of the representational techniques that make ideas, fears, etc. manifest. I will begin by outlining the special character of the image construed as the picture of something; then I will review the necessary cognitive capabilities of perception and depiction, and finally account for the role of pictorial representation in cognitive evolution.

## 1. The Riddle of Images

According to a widely accepted conception, »an image cannot be seen *as such* without the paradoxical trick of consciousness, an ability to see something as ›there' and ›not there' at the same time« (MITCHELL 1987: 17); or as Belting formulated the same idea in a rather enigmatic way, »what an image is: the *presence of an absence*« (BELTING 2005).

This paradox ascribed to images is based on the triad of the vehicle (the physical picture), the pictorial subject (the subject of pictorial representation), and the pictorial image (that represents its subject), and relates to the mostly implicit consideration of images as based on similarity or convention. First, I will attempt to show how the distinction of the elements of this triad relate to representationalism, and then focus on similarity.

Evan Thompson epitomizes the special character of picture-viewing with the help of the three distinct elements of a picture of something. As he put it,

The physical picture is present to perception, while the absent pictorial subject is re-presented, brought to presence by the pictorial image. In viewing and appreciating a picture, we are interested mainly in neither the physical picture nor the pictorial subject as such, but rather in the pictorial image that appears in the physical picture and represents the pictorial subject. (THOMPSON 2007: 288)

Thompson, with the help of the phenomenological distinction between the intentional act of presentation and re-presentation (THOMPSON 2007: 25f.), characterizes the main differences between perception, imagining, and picture viewing. According to this distinction, perceptual experience is always presentational (the subject of the experience is present), while remembering and imagination are re-presentational, since their subject is not present phenomenally. That is, re-presentational experience, though related to presentational experiences, is determined by the absence of its subject. Either we try remembering a face (and/or scenery), or by imagining it we represent it, even if we rely on earlier presentational experiences.

Even though we can, with the help of the conceptual distinction between the intentional act of presentation and re-presentation, nicely define

the distinctive features of the abovementioned activities,<sup>1</sup> this does not help us in understanding how these engagements necessitate and facilitate each other. Moreover, the given description of picture-viewing clearly performs the dualism of form and content, and therefore suggests the possibility of their separation. As James, Wittgenstein, and Merleau-Ponty<sup>2</sup> declared, we have no ground for distinguishing pain or a thought from its expression. Can we see a pictorial image without its vehicle?

Consider what it is that we recall when we try to remember Vermeer's *Girl with a Pearl Earring*: the girl (the subject of the pictorial image), the pictorial image, or the painting in its corporeality? How it is possible to identify one's intentions without the expression of them with either movements, gestures, a change of facial expression, or with words? As Arnheim quite explicitly states, »[n]either the formal pattern, nor the subject matter is the final content of the work of art. Both are instruments of artistic form. They serve to give body to an invisible universal« (ARNHEIM 1974: 461). That is, in the case of works of art, both the subject of the pictorial image and the pictorial image itself serve the purpose of expressing an idea.

The dualism of form and content is, I believe, rooted in representationalism, since it suggests from the very beginning (at least from Descartes) that representations are the mediators between the physical world and mental activity. Similarly, in the course of the imagery debate from the 1970's, descriptivism asserts that mental representations have a propositional structure, while pictorialism (the opposite pole) suggests mental representations are depictive. If we conceive mental representations either as being symbol manipulation in accordance with certain rules, or pictorial items in consciousness, the floor is open for a fallacy, namely believing that mental representations can be considered as either propositional or pictorial. In the case of pictorialism, as Thompson recalls Sartre's criticism, »an image is nothing other than a relation«. That is, »a mental image properly understood is not a content contained in consciousness; rather, it is a dynamic and relational act of consciousness« (THOMPSON 2007: 301).

Regarding mental images, many claim that these mental phenomena are pictorial in consciousness as well; but it is rather the »illusion of immanence« as Sartre coined this kind of fallacy (THOMPSON 2007: 300f.). On the basis of these phenomenological considerations, and taking into account motor skills, the enactive approach<sup>3</sup> suggests bodily skills and perception are inseparable, and thus provides an alternative for representationalism (be it

---

<sup>1</sup> For further details, see THOMPSON 2007: 287ff.

<sup>2</sup> See JAMES 1884: 189, WITTGENSTEIN 1958: 41f., and MERLEAU-PONTY 2008: 209ff.

<sup>3</sup> Enactivism has far-reaching roots in the history of philosophy (however, some of them are almost non-referred), closely relates to cognitive psychology, and interestingly, combines in a promising way some suggestions of so-called continental philosophy with the conceptual framework of the philosophy of mind. In 1991 F.J. Varela, E. Thompson and E. Rosch proposed »as a name the term enactive to emphasize the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of a variety of actions that a being in the world performs« (VARELA/THOMPSON/ROSCH 1991: 9).

either descriptive or pictorial), a promising way to bridge the gap between the world and consciousness.

I suggest considering resemblance and convention as being conflicting characteristics of pictorial representation (as compared to linguistic symbols), which is possible when we ignore the expressive medium and open the floor to the abovementioned fallacy. Most pictorialists suggest linguistic symbols are conventional, while pictures are based on similarity. But defining what similarity consists of seems to be a hard nut to crack. Kristóf Nyíri, on the basis of Gombrich's considerations, suggests »[r]esemblance *is* selective equivalence« (NYÍRI 2011: 17). This apparently apt formulation leads us to a great deal of difficulty. Granted, in the case of geometry this definition works very well because we can work with well-defined axioms and definitions and even among theoretical entities. Two triangles are similar if their corresponding angles are equal, and their corresponding sides are proportional. (Of course, one of the conditions automatically implies the other.) But in everyday life, we do not exclusively face clear-cut cases. We can recognize emoticons and pictograms on the basis of similarity; we can see similarity between a dog and its keeper, or two persons, between animate and inanimate beings, etc. Since the notion of equivalence is much more demanding, numerous questions arise: Can we give a list of equal items in each case? Or examined from another angle: can equal features provide sufficient ground for similarity? Is there a definite number of identical/equivalent features required to see two things as being similar? The list can continue, but my main concern is whether we can define equal features in the course of recognizing similarity when this process relates to presentational, representational, and motor skills. That is, I would say that since the recognition of resemblances is based not exclusively on visual capabilities, but is rather also related to motor activity and other modalities, then resemblance construed as selective equivalence may lead us astray. How can we talk about equivalence when we speak of different kinds of related capabilities? If we accept that visual perception and motor engagement are inseparable, analogy seems to be a more appropriate candidate. When we apprehend the ambient world, we rely on earlier experiences, and while utilizing them we act upon this world with the help of different motor skills.

As Harald Høffding suggested, Kant's essential invention was the introduction of analogy, *i.e.* »likeness of the relations of different phenomena, not likeness of single qualities« (HØFFDING 1905: 200). Analogy opens up a new horizon in theorizing. Although we do not gain much regarding what similarity consists of, with the introduction of analogy, as opposed to the requirement of equivalence, the difficulty raised by the demand of relating different kinds is eliminated. (I have in mind the antinomic situation when we need to talk about different things in terms of identical features.) How can a physical and a mental image have identical features when they are different in kind? I believe analogy and the exclusion of the »illusion of immanence« provide an appropriate ground for the understanding of how images work regarding

both perception and depiction. As Kant introduced analogy and underscored that a schema is not identical with an image, it became possible to anchor mental activities in the physical, flesh and blood world.

## 2. Visual Perception and Pictorial Representation

Considered through the lens of Gestalt psychology, picture-viewing and pictorial depiction are based on the same ground, though the capability of depiction requires some additional knowledge, *viz.* the knowledge of the underlying structure of the perceived effect.

As Thompson suggests, in the case of picture-viewing we focus on the pictorial image. However, with a work of art or a piece of scientific visualization, this focus is inseparable from the pictorial subject, *i.e.* the idea/intention of the artist or the mutation of the investigated organ. On the one hand, I believe the elements of the triad on the basis of the importance of the context; on the other hand, because pictorial representations are objectifications of certain thought processes, and accordingly, their appearance heavily depends on the scope and limits of their vehicle, the elements of the abovementioned triad can hardly be isolated even within the framework of theorizing. But instead of getting ahead of myself, first I will attempt to clarify the peculiarities of the apprehension of a picture of something.

Pictures are representations, or more precisely objectifications, as they are accessible to others. Since they represent something, they can be considered as symbols. The working definition of a symbol is very akin to the riddle of images. In accordance with the general view, the psychologist Judy S. DeLoache suggests »a symbol is something that someone intends to represent something other than itself« (DELOACHE 2004: 66). Many things can be a symbol, as symbolic competence evolves from »general symbolic ability«, and relates to intentionality. Young children, when focusing on a symbol itself, may encounter difficulties: it turns out that the more appealing and conspicuous an object is, the more difficult it is to consider as other than itself, and having physical contact with an object also increases the difficulty of using it as a symbol. Taking pictures as symbols into account, although pictures have a unique characteristic, *viz.* they are unavoidably perspectival<sup>4</sup>, experiments provide evidence that perspective does not play a role in the case of younger children. In one experiment, pictures of tulips taken from above and others from directly in front of the flowers were shown to children. The photos did not seem to differ considerably for children under three years of age; it appears that 3-year-old children do not consider the difference of perspective when the depicted objects are identical (DELOACHE 2004: 69).

---

<sup>4</sup> Pictures as being unavoidably perspectival should be considered in a restricted sense, *viz.* with the exception of graphs, diagrams and the like. For more details see CASATI/GIARDINO 2012: forth.

This suggests that motor skills and earlier experiences help us recognize the referent of a symbol with no regard to the perspective from which the pictorial symbol is depicted. This option creates a strange situation: according to anatomical evidence there are two visual pathways that are usually supposed to be attached to different tasks, and the motor component seems to be deeply integrated into the act of visual perception, particularly in case of picture viewing. According to the representational theory of the visual mind, one pathway is devoted to visual percepts in the service of further cognitive processing, and the other is in the service of action (cf. JACOB/JEANNEROD 2003). Though the transit between the two pathways is beyond question, the functional differences seem to provide ground for researchers and philosophers to separate the motor component from visual perception. But enactivism underscores the cooperation between the two capabilities. Valera and his co-authors suggest »sensory and motor processes, perception and action, are fundamentally inseparable in lived cognition. Indeed, the two are not contingently linked in individuals; they have also evolved together« (VARELA/THOMPSON/ROSCH 1993: 173). Noë goes a bit further when he suggests that »[t]he basis of perception [...] is implicit practical knowledge of the ways movement gives rise to changes in stimulation« (NOË 2004: 8). The anatomic distinction of dorsal (in the service of action) and ventral (dedicated to further processing of what is seen, perceptual representation, identification, etc.) paths is not to be considered either as evidence for or against the enactive approach, since enactivism proposes that any kind of perceptual representation is determined by the applied sensorimotor skills. Noë highlights the importance of motor skills even in the case of observational concepts. Experiencing something as cubical is possible because you possess some practical knowledge according to which you experience a cubic thing »as something whose appearance would vary in precise ways as you move in relation to it, or as it moves in relation to you. You have an implicit practical mastery of these patterns of change« (NOË 2004:117). Similarly, as the DeLoache experiments show, young children recognize highly realistic two-dimensional depictions of three-dimensional objects with the help of their body: they try exploring them with their hands, using them as real objects. This kind of manual exploration starts to decrease after 18 months of age (cf. DELOACHE 2004: 68). That is, the necessary condition of picture-viewing, *viz.* the recognition of a picture as the picture of something, is a capability gained thanks to bodily experiences by motor engagements in ontogeny.

The relation between perception and motor skills is fairly obvious in case of depiction. But this relation reaches beyond the requirements of the fine tuning of movements in the course of creating drawings, paintings, sculptures and the like; perception and depiction are rooted in embodied cognition. As Merleau-Ponty and current conceptual metaphor theory underscore, our concepts are based on our primordial bodily experiences.

[I]f the words ›enclose‹ and ›between‹ have a meaning for us, it is because they derive it from our experience as embodied subjects. In space itself independently of the presence

of a psycho-physical subject, there is no direction, no inside and outside. A space is »enclosed« between the sides of a cube as we are enclosed between the walls of our room. (MERLEAU-PONTY 1962: 236)

The Gestalt laws of perception<sup>5</sup> and the idea of embodiment provided the ground for revealing the relation between perception and depiction. And while both depiction and perception are rooted in the same bodily experiences, depiction necessitates the knowledge of the representational means.

Arnheim came close to the enactive approach when he considered the perception of shapes as »an eminently active occupation«. As he writes, »in looking at an object, we reach out for it. With an invisible finger we move through the space around us, go out to the distant places things are found, touch them, catch them, scan their surfaces, trace their borders, explore their texture«<sup>6</sup> (ARNHEIM 1974: 43).

He also suggests considering vision as »a creative activity of the human mind«. It is because »[v]ision deals with the raw material of experience by creating a corresponding pattern of general forms, which are applicable not only to the individual case at hand but to an indeterminate number of other similar cases as well« (ARNHEIM 1974: 46). In visual perception, all our earlier visual experiences play a part. Because we form *perceptual concepts* while perceiving the ambient world, we are able to grasp some »outstanding features of objects« and at the same time notice slight changes. Arnheim argues on the basis of the holistic nature of perception. As his example suggests

---

<sup>5</sup> Regarding visual perception, Gestalt psychology recognized some key characteristics of visual perception. There are different groupings of these laws/principles depending on nuances and time period. Attempting to recapitulate the main characteristics, emergence is one of the main principles. Emergence suggest that a thing in our environment is perceived as a whole, all at once, i.e. the recognition of a dog comes not by identifying its parts and then by the composition of these parts, but rather via the recognition of certain patterns. The generative aspect of perception is reification. This means that the experienced percept provides more spatial information than the sensory stimulus. That is, we recognize certain shapes although they have no contours like in the case of a Kanizsa triangle. Multistability of perception refers to a tendency in ambiguous perceptual situations. As in the case of Rubin's heads/vase illusion (1919) the figure (the substantial appearance of an object) and the ground (the general environment of the object) oscillate in accordance with the interpretation. Thanks to invariance, it is possible to recognize simple geometrical objects independently from the perspective we see them in and elastic deformations. According to the law of Prägnanz, we tend to arrange our experience in a manner that is simple, symmetric, regular and orderly. There are six principles of Gestalt perception: proximity, similarity, common fate, good continuation, closure, and symmetry. The principle of proximity/contiguity state that objects closer to each other will be seen as ones belonging together. The principle of similarity asserts that objects which share visible characteristics (such as shape, colour, size, orientation, etc.) will be seen as belonging together. The principle of common fate/uniform destiny captures the idea that movement or function in a similar direction or manner will be perceived as a unit. The principle of good continuation predicts the preference of continuous figures (in case of crossed lines we perceive two lines crossing instead of four meeting at a point). According to good closure we tend to see complete figures despite any missing information, i.e. we tend to organize the visual information to make a whole. The principle of symmetry relates to the figure-ground relationship, as it states that when we perceive objects we tend to perceive them as symmetrical shapes that form around their centre. Thus, symmetrical images are perceived collectively even in the distance and symmetrical areas tend to be seen as figures against the background of asymmetrical background.

<sup>6</sup> Cf. »My theme ... is the touch-like character of vision itself, by which I mean precisely the fact that vision too is intrinsically active« (NOË 2004: 97).



The minute modifications of muscle tension or skin color that make a face tired or alarmed are quickly observed. The viewer may, however, be unable to pinpoint what has caused the change in overall appearance because the telltale signs fit smoothly into an integrated framework. (ARNHEIM 1974: 44)

That is, we perceive general patterns deeply integrated into the given context, therefore even slight changes can modify the overall structure, and therefore they are easily noticeable. Arnheim distinguishes perceptual and representational concepts. The latter are necessary when in a particular medium the relevant features of a model are to be represented.

[I]mage-making of any kind requires the use of representational concepts. Representational concepts furnish the equivalent, in a particular medium, of the visual concepts one wishes to depict, and they find their external manifestation in the work of the pencil, the brush, the chisel. The formation of representational concepts, more than anything else, distinguishes the artist from the nonartist. (ARNHEIM 1974: 169)

As we can see, representational concepts are rooted in perceptive and motor skills and, importantly, in the recognition of the underlying structure of the perceivable effect on the one hand, and the knowledge of the possibilities and limits of the given expressive medium on the other hand. Young children are capable of recognizing the difference between vertical/horizontal and oblique lines, but only later are able to actually produce the latter. Children's drawings reveal much about the relation of perceptive, depictive and motor skills. Though a straight line is visually the simplest line, the first scribbles are whirl-like »circular path of lines«. Arnheim suggests two reasons for this interesting phenomenon. On the one hand, »the child's arm re-enacted the motion as a gesture during the drawing« (ARNHEIM 1974: 172), and on the other hand, according to the law of simplicity, a circle is the »simplest visual pattern« as it »does not single out any one direction« (ARNHEIM 1974: 175) and moreover, certain anatomic features subserve curved motion.<sup>7</sup> Experiments suggest the young children favour circles over other shapes (ARNHEIM 1974: 175), and early childhood drawings tend to show humans as round tadpole-like shapes. Of course, these scribbles are not circles in a strict sense, but the result of the child's more or less uncontrolled movements is most similar to circles.

Recent research in the field of developmental psychology has provided evidence that the intention of communication plays an important role in drawings in ontogeny. »The first scribbles of a child are not intended as representation. They are a form of the enjoyable motor activity in which the child exercises his limbs, with the added pleasure of having visible traces« (ARNHEIM 1974: 171). With practice, the schemas become more detailed, and details gain importance when it turns out the drawing can be considered as a representation of something/someone. Four-year-old children do not take into account details that can help when it is important to distinguish a balloon and a lollipop. But

---

<sup>7</sup> Arnheim mentions the »lever construction of the human limbs« and the structure of the shoulder joint. The movement by the elbow, the wrist, and the fingers also facilitate rotation. (ARNHEIM 1974: 174f.)

[t]hree- and four-year-olds produced better drawings when they understood that their drawings would be used to communicate to an adult which object to use in a game than when they were just asked to draw the objects. When informed that their drawings had not communicated clearly enough, children of this age subsequently improved them. Thus, having a communicative intent enhanced the children's appreciation of the symbolic function of their drawings. (DELOACHE 2004: 67f.)

The idea of embodiment provides a frame of reference in which perception and depiction are inseparably attached to motor skills; and depictive capability relates to social commitments and necessitates material engagement as it enables expression within a given expressive medium. The necessity of considering the other as an intentional agent, the intention of communication, *i.e.* beyond the joyful activity of drawing, being aimed at sharing some ideas is a necessary condition of pictorial representation in ontogeny.

### 3. Representations and Material Engagement

The view that sociability is a fundamental characteristic of human nature is widely accepted in the fields of sociology, linguistics, cognitive evolution theories, developmental psychology, etc. In his social brain hypothesis, anthropologist and evolutionary psychologists Robin Dunbar suggests that the size of the neocortex and that of the group correlate.

The social brain hypothesis implies that constraints on group size arise from the information- processing capacity of the primate brain, and that the neocortex plays a major role in this. ...The constraint on group size could be a result of the ability to recognize and interpret visual signals for identifying either individuals or their behavior; limitations on memory for faces; the ability to remember who has a relationship with whom (e.g., all dyadic relationships within the group as a whole); the ability to manipulate information about a set of relationships; and the capacity to process emotional information, particularly with respect to recognizing and acting on cues to other animals' emotional states. (DUNBAR 1998: 184)

That is, taking part in a group's life requires more than just the capability of remembering, it requires additional capacities of information processing. Merlin Donald, in his cognitive evolutionary theory, relies on Dunbar's hypothesis (expressed in earlier writings by him) and goes a bit further when he underscores that new institutions/inventions can extend our brain capacity. He suggests »that the evolution of cognitive structure *at the modular level* might have continued well beyond the point at which physical evolution had stopped«. Cultural patterns have a far-reaching effect on cognition, even »in terms of its fundamental neurological organization« (DONALD 1991: 14). He differentiates three main transitions in phylogeny: a shift from episodic to mimetic, from mimetic to mythical, and finally from mythical to theoretical culture. Each of these changes is a consequence of the emergence of a new

kind of representational skill and there is a concomitant increasing load on biological memory, with the exception of the last change.<sup>8</sup>

Mimetic representation as a supplement to episodic memory yielded the ground for creating a community with special habits and organization, conferring a sense of identity on the members of community. The emergence of linguistic representation, (beside episodic memory and mimetic skills), made the evolution of a framework possible in which mythical constructions could emerge as structuring and orienting potentials within the communities. The transition from mythical to theoretical culture was thanks to a new invention, *viz.* the *external symbolic storage system*. While earlier changes inserted a huge load on memory capacity, this latter one rather exempted biological memory. This apparently technological invention established considerable alterations in cognitive habits.

[T]heoretic culture broke with the metaphoric style of meaning in oral-mythic culture. Where narrative and myth attribute *significances*, theory is not concerned with significance in the same sense at all. Rather than modeling events by infusing them with meaning and linking them by analogy, theory dissects, analyzes, states laws and formulas, establishes principles and taxonomies, and determines procedures for the verification and analysis of information. It depends for its advanced development on specialized memory devices, languages, and grammars. (DONALD 1991: 274f.)

According to the Donaldian chronology, pictorial representations emerged within the framework of oral-mythical culture and gain importance in Donald's theory because of their mediating role in the development of external storage; images are considered as the first steps towards external storage. As he put it »the first pictorial images themselves were also external representations. They existed outside of the individual, rather than in visual memory. Therefore a technological bridge was under construction that would eventually connect the biological individual with an external memory architecture« (DONALD 1991: 284).

Donald considered cognitive evolution as being an encapsulating development of representational skills. The refinement of representational skills provides the ground for the evolution of increasingly complex communities, and this entails further representational capabilities. Accordingly, Donald suggests that while the invention of pictorial representation presupposes earlier cognitive skills, »pictorial images, on one level, must possess episodic reference«, and »mimetic manufacturing skill and linguistic conceptual skill« (DONALD 1991: 283), pictorial images also paved the way for the creation of an external and lasting form of representation. Therefore, these first steps opened up the horizon of cognitive development. An encapsulating evolution of representational skills towards external storage implies, as Donald sug-

---

<sup>8</sup> »Human memory had, from its inception, expanded the range of primate memory. The earliest form of hominid culture, mimetic culture, depended on an expansion in the self-representational systems of the brain and created the initial base for semantic memory storage, which consisted initially of representational action scenarios reflected in mime, gesture, craft, and skill. With the evolution of speech and narrative ability, there were even greater increases in the load on biological memory, adding not only the storage networks for phonological rules and the lexicon in its entirety but also a very large store of narrative conceptual knowledge« (DONALD 1991: 319).

gests, that »we are permanently wedded to our great invention, in a cognitive symbiosis unique in nature« (DONALD 1991: 356). Thus, human minds can be considered as being hybridizations.

At this point, I will outline the view of the enactive approach with regard to the role of images in phylogeny. Though this view calls into question the necessity of linguistic conceptual skills, I will underscore the importance of external representations in both cases.

The alternative view suggests emphasizing the importance of material engagement as compared with Donald's account of the evolution of the human mind. It highlights the importance of the interaction between the cognizing agent and its environment. Under the heading of neuroarchaeology, we can find recent efforts towards converging the results of neuroscience and the findings of archaeology as it relates to material culture. As Lambros Malafouris put it,

neuroarchaeology is strictly an interactionist approach, aiming primarily to understand the bidirectional links between brains, minds and culture. Brain activity is a crucial component of the human mind but so is also material culture. The human mind as an interactive, embodied, and distributed autozoetic system is neither restricted to nor can it be identified with the temporal and spatial boundaries of the brain within. (MALAFOURIS 2010: 64)

Against the background of the notions of embodiment, scaffolding, and material engagement, pictorial images enter a new light. Within the framework of embodied cognition, the importance of the ambient world and cooperation with it radically increases. The responsive relation suggests a symbiosis between the biological brain, the body, and its environment. According to the extended mind theory and also Donald himself, the ambient world provides a scaffold, or viewed the other way around, in order to accommodate to their conditions, humans (and other animals) rely on affordances, *i.e.* the easing elements of the environment. This kind of accommodation drives changes in the organism as well. Accordingly, the capability of pictorial representation as a kind of material engagement can be considered as a new way of accessing the world. As Malafouris suggests,

[t]hrough the process of »imaging« the underlying mechanisms of human perception are being transformed to an object *for* perception and contemplation. Those invisible mechanisms become now perceivable visual patterns arrayed and combined in real time and space. In this sense, the image offers a new mode of epistemic access to the world of visual experience. (MALAFOURIS 2007: 299)

The main difference between the Donaldian approach and Malafouris' suggestion lies in the relation between pictorial skills and conceptual capacity. Malafouris

believe[s] that to say that a painting from Chauvet »resembles«, or »looks like«, a rhinoceros, a horse or a bull, does not necessarily imply that these paintings also »represent« those animals in a concept-mediated referential manner. Iconicity does not in itself imply the existence of consciously manipulable content-bearing tokens, though it certainly sets up [...] a visually stimulating channel of influence for their creation. (MALAFOURIS 2007: 293)

Unlike Malafouris, Donald seems to believe that there is an underlying »linguistic conceptual skill« (DONALD 1991: 283). This difference,<sup>9</sup> though not negligible, does not play an important role for our current purposes. Under both frameworks, it is possible to develop a theory of images as they inherently belong to the cognitive capacity of human beings. Beyond the striking similarities between the prerequisites of linguistic<sup>10</sup> and pictorial capabilities, scaffolding and new expressive forms entail considering images as an important anthropological issue. Nevertheless, from an enactive point of view, images definitely play a crucial role in phylogeny.

The Palaeolithic image-maker constructs an external scaffold that affords the world to be seen and experienced in ways that the physiology of the naked eye by itself does not allow. This scaffolding also enables a new direct understanding of the human perceptual system and thus offers to the Palaeolithic individual the opportunity to become in some sense, maybe for the first time, the engineer of his or her own perception. The image, as it is also the case with language, enabled humans *to think about thinking*. (MALAFOURIS 2007: 300)

#### 4. Conclusion

Being able to create a picture as a picture of something and using linguistic symbols require the same conditions: a social brain as a prerequisite, and all this entails: a community, representational skills, and the integration of new representational patterns into everyday practice. Representations as means of communication are bounded to habits, institutions (in a community) and/or some kind of material support. Casting a glance at ontogeny, we can see that symbol use, picture viewing, linguistic and depictive skills are the result of being a member of a group (family, nursery, etc.) and thus becoming involved in its communicational habits.

In ontogeny, the manual exploration of realistic images is a prerequisite of picture-viewing in a broad sense. The capability of pictorial depiction needs an intention of communication and perfection in the expressive medium. Accordingly, image production presupposes sociability. Since many scholars share the view that the social brain is the answer to the puzzling question of how human cognition could reach its recent capacity in a relatively short time period in the course of biological evolution, it can be considered as being particularly human. Having a social brain necessitates communica-

---

<sup>9</sup> I believe this difference is rooted in the broader context of the two investigations: Donald aimed to create continuity in cognitive evolution focusing on representational skills. Malafouris takes into account cave paintings from a different perspective, viz. from an enactive approach. But both scholars agree regarding the long-lasting effect of external representations. This is unavoidable because, as I suggested, the core idea of the extended mind theory can be found in Donald's 1991 book (cf. KONDOR 2008: 77ff.).

<sup>10</sup> Michael Tomasello (1999) suggests cultural transmission is a decisive element in the evolution of the cognitive setting of the modern human mind. Accordingly, intersubjectivity (as a precondition), communication, and therefore linguistic capability play a crucial role. Regarding linguistic skill, he underscores the importance of the possibility of external expression as it facilitates cognition (cf. TOMASELLO 1999: 195). The perspectival character of language (namely, we express ourselves in a way that best fits our purposes) and the representational concept (as Arnheim suggested) similarly presuppose expertise in the given expressive medium.

tion. Communication presupposes the capability of objectification, *i.e.* sharing an experience/idea is possible only if it is accessible for others. Accordingly, representation plays a crucial role. Since representation is only possible as presentation, the expressive medium is decisive.

Representationalism as a highly resistant view in philosophy entered the scene of philosophical debate with the Cartesian dualism of the extended body and the immaterial mind quite explicitly. Under this dualist framework, we do not have a satisfying answer to the question of how it is possible to make a physical percept accessible for mental processing. The enactive approach seems to pave the way for an answer beyond the horizon of representationalism. It regards perception as an active engagement, as opposed to a passive acceptance of external stimuli for further cognitive processing that results in an internal representation.

The inclusion of bodily skills, and hence the importance of the responsive relation between the cognizing subject and her environment, shed a new light on the act of picture viewing, pictorial depiction and, the role of pictorial representation in phylogeny. Similarly, recent psychological experiments illuminate the special role of pictures in ontogeny: we learn the difference between a picture and its subject with the help of bodily exploration, and we gradually learn how we can use an expressive medium through the joyful activity of leaving traces, and with finer coordination, how we can use this medium to express something and develop *representational concepts*. All these result in a permanent refinement of cognitive capacity. Images as external representations enter a new light: as Donald suggests, they pave the way for creating an extra-cranial storage system, and according to Malafouris, like language, they open the floor for reflective thinking within the framework of material engagement.

## References

- ARNHEIM, RUDOLF: *Art and Visual Perception. A Psychology of the Creative Eye*. Berkeley [U of California P] 1974
- BELTING, HANS: Toward an Anthropology of the Image. In: WESTERMANN, MARIËT (ed.): *Anthropologies of Art*. New Haven [Yale UP] 2005
- CASATI, ROBERTO; VALERIA GIARDINO: Public Representations and Indeterminacies of Perspectival Content. In: KONDOR, ZSUZSANNA (ed.): *Enacting Images. Representation Revisited*. Cologne [Halem] 2012, forthcoming
- DELOACHE, JUDY S.: Becoming Symbol-Minded. In: *Trends in Cognitive Sciences*, 8(2), 2004, pp. 66–70
- DONALD, MERLIN: *Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA [Harvard UP] 1991

- GOMBRICH, ERNST: *Art and Illusion: A Study in the Psychology of Pictorial Representation*. London [Phaidon Press] 1960
- DUNBAR, ROBIN I.M.: The Social Brain Hypothesis. In: *Evolutionary Anthropology*, 6(5), 1998, pp. 178–190
- HØFFDING, HARALD: On Analogy and Its Philosophical Importance. In: *Mind*, 14(54), 1905, pp. 199–209
- JACOB, PIERRE; MARC JEANNEROD: *Ways of Seeing. The Scope and Limits of Visual Cognition*. Oxford [Oxford UP] 2003
- JAMES, WILLIAM: What is Emotion? In: *Mind*, 9(34), 1884, pp. 188–205
- KONDOR, ZSUZSANNA: Being Mobile. Cognitive Multiplicity. In: KONDOR, ZSUZSANNA: *Embedded Thinking. Multimedia and the New Rationality*. Frankfurt/M. [Peter Lang] 2008
- MALAFOURIS, LAMBROS: Before and Beyond Representation. Towards an Enactive Conception of the Palaeolithic Image. In: RENFREW, COLIN; IAIN MORLEY (eds.): *A Global Prehistory of Figurative Representation*. Cambridge [McDonald Institute of Archaeological Research] 2007, pp. 287–302
- MALAFOURIS, LAMBROS: Metaplasticity and the Human Becoming. Principles of Neuroarchaeology. In: *Journal of Anthropological Sciences*, 88, 2010, pp. 49–72
- MERLEAU-PONTY, MAURICE: *Phenomenology of Perception*. Translated by Colin Smith. London [Routledge] 1962
- MITCHELL, WILLIAM J. THOMAS: *Iconology. Image, Text, Ideology*. Chicago [U of Chicago P] 1987
- NOË, ALVA: *Action in Perception*. Cambridge, MA [MIT Press] 2004
- NYÍRI, KRISTÓF: Gombrich on Image and Time. In: SACHS-HOMBACH, KLAUS; RAINER TOTZKE (eds.): *Bilder – Sehen – Denken*. Cologne [Halem] 2011
- THOMPSON, EVAN: *Mind in Life. Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA [The Belknap Press of Harvard UP] 2007
- TOMASELLO, MICHAEL: *The Cultural Origins of Human Cognition*. Cambridge, MA [Harvard UP] 1999
- VARELA, FRANCISCO J.; EVAN THOMPSON; ELEANOR ROSCH: *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge [MIT Press] 1993
- WITTGENSTEIN, LUDWIG: *The Blue and Brown Books*. Oxford [Basil Blackwell] 1958

Jakob Steinbrenner

## **Was heißt Bildkompetenz? Oder Bemerkungen zu Dominic Lopes' Kompetenzbedingung**

### **Abstracts**

In his book *Understanding Pictures* (1996), Dominic Lopes suggests that a picture theory should be checked for four conditions. One of them is the condition of competence. It remains unclear whether these four conditions have a specific status or precise validity for specific persons. It is a goal of my paper to resolve or to reduce these uncertainties. In order to do that, I start sketching the motivation and the method of Lopes in part one. Then I present the condition of competence in part two and explain its function in the third part. This leads to the question whether the condition of competence can be applied only to pictures and picture systems or also to terms and languages. Another question behind my reflections is: Why picture philosophy? My explanations however provide an implicit answer to that question at best.

Dominic Lopes schlägt in seinem Buch *Understanding Pictures* (1996) vier Bedingungen vor, die Bildtheorien erfüllen sollen. Eine davon ist die Kompetenzbedingung. Unklar bleibt, für wen diese und die anderen drei Bedingungen gelten sollen und welchen Status sie haben. Diese Unklarheiten zu beseitigen oder zumindest zu vermindern, ist ein Ziel meiner Überlegungen. Dazu werde ich im ersten Teil Lopes' Motivation und Methode kurz beschreiben, im zweiten Teil die Kompetenzbedingung vorstellen und im dritten Teil darauf eingehen, wozu sie dient. Dies wird zur Frage führen, ob sie allein auf Bilder beziehungsweise Bildsysteme oder auch auf Ausdrücke beziehungsweise Sprachen anwendbar ist. Eine weitere Frage, die hinter meinen Überlegungen



steht, lautet: Wozu Bildphilosophie? Meine Darlegungen werden bestenfalls eine implizite Antwort auf diese Frage geben.

## 1. Lopes' Motivation und Methode

Lopes geht in seinem Buch *Understanding Pictures* der Frage nach, wie es Bildern möglich ist, uns unter geeigneten Umständen Erfahrungen zu vermitteln, die der Welt Eigenschaften zuschreiben, wie sie durch Bilder repräsentiert werden:

(T1) Wenn wir ein Bild auf die richtige Weise erfahren, haben wir eine Erfahrung, welche die Welt mit Eigenschaften repräsentiert, die das Bild der Welt zuschreibt (vgl. LOPES 1996: 4).

Dafür, dass eine Bildtheorie T1 erklären kann, muss sie nach Lopes vier Bedingungen genügen. Warum es genau vier Bedingungen und nicht noch weitere sind, dazu erfahren wir wenig. Man muss sich darüber im Klaren sein, dass diese Bedingungen in einer Metatheorie zweiter Stufe formuliert sind, d.h., sie sollen für Bildtheorien gelten, die den Anspruch haben, T1 zu erklären. T1 selbst kann dabei sowohl als Satz der Umgangssprache wie auch des metatheoretischen Ansatzes von Lopes verstanden werden. Was für einen Status hat T1 in der Umgangssprache? Ist es beispielsweise eine umgangssprachliche Erklärung wie folgende?

›Max hat Nasenbluten, weil er einen Schlag auf die Nase erhielt.‹

Hiergegen lässt sich einwenden, dass es in T1 nicht um die Erklärung einer Tatsache geht, sondern um die Erklärung, wie ein Ausdruck gebraucht wird. Viele Philosophen sprechen in diesem Fall von einer begrifflichen Erklärung. Ich ziehe an dieser Stelle den Ausdruck ›grammatikalischen Satz‹ vor, wie er von Wittgenstein verwendet wird, da der Ausdruck ›Begriff‹ in philosophischen Kontexten für gewöhnlich einen ganzen Rattenschwanz von Problemen mit sich bringt.

Grammatikalische Sätze zeichnen sich im Gegensatz zu gewöhnlichen beschreibenden Sätzen dadurch aus, dass sie festlegen, wie bestimmte Ausdrücke zu verwenden sind. Auf T1 bezogen heißt das, T1 dient dazu, den Gebrauch des Ausdrucks ›ein Bild auf die richtige Weise erfahren‹ zu bestimmen. Wichtig für die weiteren Überlegungen ist, dass ein grammatikalischer Satz Ausgangspunkt einer metatheoretischen Überlegung zweiter Stufe ist. Für Lopes bedeutet dies, von jeder brauchbaren Bildtheorie, die T1 erklären will, zu fordern, dass sie seinen vier Bedingungen genügt.<sup>1</sup>

---

<sup>1</sup> Eine Frage dabei ist, ob T1 in der Umgangssprache und der Metatheorie dieselbe Bedeutung hat. Eine Antwort darauf setzt offensichtlich eine Klärung des Ausdrucks ›bedeutungsgleich‹ voraus – bekanntermaßen ein philosophisches Mammutprojekt.

Welche Theorien hat Lopes dabei im Blick? Es sind philosophische Theorien von Wollheim, Walton, Peacocke, Goodman und anderen (vgl. STEINBRENNER 2009). Wollen diese aber T1 oder überhaupt etwas erklären? Erinnerung sei an Wittgensteins Ausspruch, alle Erklärung muss weg und an ihre Stelle bloße Beschreibung treten. Für Wittgenstein hieß dies, Philosophie kann einzig die Grammatik verschiedener Sprachspiele beschreiben und sonst nichts.<sup>2</sup> Eine solche rigide philosophische Auffassung findet sich in der Bildphilosophie nicht. Ich halte das für bedauerlich, da eine ausführlichere philosophische Untersuchung der Grammatik unserer Rede über Bilder bis heute fehlt.

Die Theoretiker, die Lopes im Auge hat, fühlen sich nur begrenzt dem Wittgensteinschen Programm verpflichtet. Was nun ist ihr Ziel und ihre Methode? Formulieren wir es schwach: Alle sind daran interessiert, zu verstehen, wie es dazu kommt, dass Bilder für uns Gegenstände, Ereignisse etc. darstellen können. Sie untersuchen dazu nicht oder nur begrenzt die Grammatik, sondern entwickeln Thesen zur Struktur von Bildern, dem Wahrnehmen und Verstehen von Bildern, der Kommunikation durch Bilder etc. Was für einen Status haben aber diese Thesen? Es sind sicherlich keine empirischen Thesen, auch wenn sie häufig auf Thesen aus empirischen Theorien aufbauen. Gleichwohl wird mit ihnen ein allgemeiner Anspruch vertreten, insoweit sie dazu dienen sollen, prinzipielle Strukturen offenzulegen. Dazu wird in diesen Theorien auf Untersuchungen aus anderen Bereichen der Philosophie – wie etwa der Erkenntnistheorie, Ontologie, Sprachphilosophie – zu unserer Rede über Bilder und einem begrenzten Bildmaterial zurückgegriffen. Hieraus wird in den in Frage stehenden Theorien eine Fachterminologie entwickelt, mit deren Hilfe bestimmte grundsätzliche Fragen geklärt werden sollen. Wie aber stehen diese Theorien zu dem Wittgensteinschen Programm? Ich denke, die analytischen bildphilosophischen Theorieansätze, auf die Lopes zielt, werden zum einen dazu entwickelt, unseren lebensweltlichen Umgang mit Bildern besser zu verstehen, wobei ihre Anhänger glauben, dass hierzu eine bloße Beschreibung der Sprache im Sinne Wittgensteins nicht ausreicht. Zum anderen handelt es sich bei diesen Theorien um Metatheorien, die Theorien untersuchen und bewerten, deren Thema *Bilder* sind. Analytische Bildtheoretiker verwenden dazu eine Sprache, für die sie den Anspruch erheben/von der sie meinen, dass sie relativ nah an der Umgangssprache und zuweilen bestimmter Wissenschaftssprachen ist. Ein Grund hierfür ist, dass sie wünschen und glauben, damit eine relativ klare und verständliche Sprache zur Verfügung zu haben. Für orthodoxe Wittgensteinianer lässt sich dieser Wunsch aber nicht verwirklichen und dies aus zwei Gründen. Erstens, analytische Bildtheorien dieses Zuschnitts sind Pseudotheorien, da sie keine empirischen Hypothesen aufstellen, und zweitens, Philosophie als abschließende Metasprache muss sich immer der Umgangssprache bedienen, der Sprache, in der wir schlussendlich immer Theorien beschreiben und beurteilen.

---

<sup>2</sup> Begleiterscheinung dabei kann sein, der berühmten philosophischen Fliege den Weg aus dem Fliegenglas zu zeigen.

Zu einem gewissen Grade teilt Lopes vermutlich diese Auffassung. Für ihn sind seine vier Bedingungen philosophische und damit metatheoretische Bedingungen, die er aus unserem praktischen Umgang und damit unserer Rede über Bilder zu gewinnen glaubt. Sie sollen ihm dazu dienen, die Schwächen von Ähnlichkeits- und Symboltheorien und den Vorteil seiner eigenen Aspekt-Wiedererkennungstheorie aufzuzeigen. Von zentraler Bedeutung ist dabei die Kompetenzbedingung, da sie in der Auseinandersetzung zwischen Ähnlichkeits- und Zeichentheoretikern eine wichtige Rolle spielt.

## 2. Lopes' Kompetenzbedingung

Die Kompetenzbedingung setzt sich für Lopes aus zwei Komponenten zusammen:

- **Generativitätsbedingung:** Eine Bildtheorie muss erklären können, wie es dazu kommt, dass unbekannte Bilder nicht auf die Weise erlernt werden müssen wie unbekannte Wörter oder Ausdrücke. Wenn wir einmal ein bildliches Darstellungssystem verstanden haben, dann können wir jedes weitere Bild, das in dieser Weise hergestellt ist, verstehen, vorausgesetzt das Sujet ist uns vertraut. In diesem Sinne sind Bilder generativ (vgl. LOPES 1996: 70).
- **Transferenzbedingung:** Wenn jemand mit einem bildlichen Darstellungssystem vertraut ist, dann kann er mit Hilfe von Bildern Dinge, die ihm bis dato unbekannt sind, in der Welt erkennen. Das heißt, er kann sein Bildwissen transferieren (vgl. LOPES 1996: 71).<sup>3</sup>

Während die Generativität also voraussetzt, dass die Sujets bekannt sind, erlaubt die Transferenz, neue Sujets der Bilder zu erlernen. Ähnlichkeitstheorien besitzen nun offenbar genau die Stärke, diese beiden Fähigkeiten zu erklären. So ist es beispielsweise die Umrisslinie, (vgl. HOPKINS 1998; HYMAN 2006) die Bild und dargestellter Gegenstand teilen, die dem Bildbetrachter erlaubt, das Motiv zu erkennen.

Eine Erklärung in dieser schlichten Form steht dem Zeichentheoretiker nicht zur Verfügung; ihm zufolge sind Bilder arbiträre Zeichen, deren Bedeutung wir ähnlich Sprachen erlernen müssen. Die Arbitrarität der Bilder widerspricht daher dem ersten Anschein der Fähigkeiten zur Transferenz und Generativität im Umgang mit Bildern.

Bevor ich auf diese Diskrepanz zwischen Ähnlichkeits- und Zeichentheorie genauer eingehe, sei auf einige grammatikalische Sätze verwiesen, mit denen sich die beiden Komponenten der Kompetenzbedingung motivieren lassen. Ein Problem ist der Ausdruck des Darstellungssystems. Dieser Begriff ist offensichtlich kein umgangssprachlicher und zudem bleibt in der

---

<sup>3</sup> Den Ausdruck ›transfer‹ in diesem Zusammenhang hat erstmals SCHIER 1986: Kap. 3 verwendet.

ganzen bildwissenschaftlichen Literatur unklar, wie sich Bildsysteme individualisieren lassen.

Ungeachtet dessen können wir festhalten, dass Menschen, die mit bestimmten ›Arten‹ von darstellenden Bildern vertraut sind, zumeist jedes neue Motiv, soweit es ihnen aus der realen Welt bekannt ist, sogleich erkennen können. So folgt aus der Annahme, dass eine Person mit einem Bildsystem vertraut ist, also bisher in den meisten Fällen sogleich erkannt hat, was ein Bild aus dem betreffenden System darstellt, auch neue Motive erkennt, soweit ihr die Gegenstände aus der Erfahrung bekannt sind.<sup>4</sup> In diesem Sinn lässt sich der Satz (G) als grammatikalischer verstehen:

- (G) Wenn eine Person P mit einem darstellenden Bildsystem B\* vertraut ist, dann kann P im Allgemeinen sogleich neue Motive des Bildsystems B\* erkennen, insofern ihr die dargestellten Gegenstände etc. bekannt sind.

Aus Vertrautheit mit einem Bildsystem folgt daher im Allgemeinen, dass die Person neue Motive erkennt. Wir können also beispielsweise aus dem Satz

- (1) Max kennt sich mit dem Bildsystem B\* aus, auf Satz (2) schließen:  
(2) Max kann neue Motive in B\* erkennen, insofern ihm die dargestellten Gegenstände etc. bekannt sind.

Auf ähnliche Weise wie die Generativitätsbedingung lässt sich die Transferenzbedingung erklären.

- (T) Wenn eine Person mit einem darstellenden Bildsystem B\* vertraut ist, dann kann P im Allgemeinen ihr bis dato unbekannte Gegenstände etc. aufgrund eines Bildes aus B\* sogleich erkennen.

Fügt man die Bedingungen (G) und (T) zusammen und verstärkt sie im Sinne einer Äquivalenz erhält man (K):

- (K) Eine Person P ist mit einem darstellenden Bildsystem B\* genau dann vertraut, wenn P im Allgemeinen (i) sogleich neue Motive des Bildsystems B\* erkennen kann, insofern ihr die dargestellten Gegenstände etc. bekannt sind, und (ii) sie ihr bis dato unbekannte Gegenstände etc. aufgrund eines Bildes aus B\* sogleich erkennen kann.

Gegen (K) kann eingewendet werden, dass – wie schon erwähnt – unklar ist, was ein Bildsystem oder eine Art von Bildern ist und die Folgerungsbeziehungen in einer oder beiden Richtungen zu stark sind. Wenn der Einwand

---

<sup>4</sup> Damit soll natürlich nicht bestritten werden, dass eigentlich bekannte Gegenstände, wenn sie in ungewohnten Perspektiven, Beleuchtungen etc. dargestellt werden, nicht erkannt werden.

zutreffend ist, folgt daraus, dass es sich um keinen grammatikalischen Satz im Sinne Wittgensteins handeln kann. Grundlage für den Dissens über den Status von K sind also unterschiedliche Sprachintuitionen. Diese können zu folgenden beiden Einwänden führen:

- (a) K ist falsch.
- (b) K ist nicht falsch, aber aufgrund des Dummys ›Bildsystem‹ nichtssagend.

Der Einwand (a) beruht auf der Annahme, dass aus einer Vertrautheit mit einem Bildsystem nicht (i) und (ii) folgt und aus (i) oder (ii) keine Vertrautheit mit einem Bildsystem folgt. Im einen Fall wird der Einwand auf einem klaren und unstrittigen Gegenbeispiel beruhen, im anderen Fall nur auf unterschiedlichen Sprachintuitionen. Der Einwand (b) beruht dagegen auf der Annahme, dass der Ausdruck ›Bildsystem‹ keiner der natürlichen Sprache ist und damit keine Gebrauchsregeln für ihn vorliegen. Daher ist er gewissermaßen ein Dummy. Wenn die Einwände (a) oder (b) zutreffend sind, wie kommt es dann dazu, dass wir zumindest das Gefühl haben, (K) zu verstehen? Der Kritiker von K kann hier zu zwei Erklärungen geben:

- (b1) Mit K wird versucht, einen neuen Ausdruck (›Bildsystem‹) in die Umgangssprache einzuführen, d.h. eine neue Regel. Fraglich ist es, ob auf diese Weise ein neuer Ausdruck in die Umgangssprache einfach eingeführt werden kann.
- (b2) K ist ein Axiom oder eine Definition einer Theorie oder zumindest eines Theorieansatzes. Ziel der Theorie ist es, Klarheit in unsere Rede über Bilder zu bringen. Ob dies aber gelingt, ist äußerst fraglich, da die Ausdrücke ›Bild‹ etc. in der Theorie nicht mehr dieselbe Bedeutung wie in der ursprünglichen Sprache besitzen.

Meines Erachtens kann ein Vertreter von K mit dem Einwand (b1) leben. Ob der Versuch gelingen wird, den neuen Ausdruck ›Bildsystem‹ einzuführen, wird die Zukunft zeigen. Im schlimmsten Fall ist es ein folgenloser Fehlschlag. Der Einwand (b2) ist dagegen fundamentaler Art. Er beruht auf einem holistischen Sprachmodell, das bekanntermaßen bestimmte Stärken hat. Gleichwohl lässt sich an dieser Stelle dagegen argumentieren. So lässt sich die These vertreten, dass die meisten syntaktisch identischen Ausdrücke und Sätze in der Theorie wie auch in der Umgangssprache dieselbe Bedeutung haben. Sonst wäre eine Theorie, in der neue Ausdrücke eingeführt werden, unverständlich. Gleiches kann auch für Theorien zutreffen, die K erhalten. Daher ist der Einwand (b2) nicht zwingend.

Aber auch wenn man das holistische Modell akzeptiert, folgt daraus nicht notwendigerweise, dass die Theorie nicht Einblicke in die Rede über Bilder ermöglichen kann. So gilt ja für Theorien allgemein, dass sie auf der Einführung neuer Ausdrücke oder der Veränderung der Verwendung von

Ausdrücken der Umgangssprache beruhen. Eine Theorie, die K erhält, ist gleichwohl keine empirisch naturalistische Theorie, sondern eine, deren Ziel es ist, Klarheit über die Strukturen unserer Sprache zu gewinnen. An dieser Stelle ist es sinnvoll, auf eine kleine sprachliche Feinheit hinzuweisen. K ist ein Satz einer Theorie. Eine Theorie, die K enthält, erfüllt die Kompetenzbedingung, weil in ihr K gültig ist, aber sie bietet keine empirische Erklärung für K. Sie legt nur fest, dass bestimmte Schlüsse auf Grund des Schemas K gültig sind. So folgt beispielsweise aus Satz (1) Satz (2):

- (1) Max ist mit einem darstellenden Bildsystem  $B^*$  vertraut.
- (2) Max kann im Allgemeinen (i) sogleich neue Motive des Bildsystems  $B^*$  erkennen, insofern ihm die dargestellten Gegenstände etc. bekannt sind, und (ii) er kann ihm bis dato unbekannte Gegenstände etc. aufgrund eines Bildes aus  $B^*$  sogleich erkennen.

Soweit zum Schema oder Satz K. Was hat der aber mit der Kompetenzbedingung zu tun beziehungsweise was soll diese genauer besagen?

## 2.1 Wozu die Kompetenzbedingung?

Soll die Kompetenzbedingung besagen:

- (A) eine brauchbare empirische Theorie soll den Satz K erklären oder zumindest keine Ergebnisse liefern, die mit ihm unvereinbar sind, oder
- (B) eine philosophische Theorie über Bilder muss den Satz K enthalten?

Versteht man die Kompetenzbedingung in Lesart (B), wird klar, dass die Fragen ›Warum gelingt es Max, wenn er mit dem Bildsystem vertraut ist, neue Motive zu erkennen?‹ und ›Warum ist Max, wenn er neue Motive erkennt, mit dem Bildsystem vertraut?‹ einfach und ohne jeden Rückgriff auf eine empirische Theorie in einer philosophischen Theorie zu beantworten sind. Grund hierfür ist, dass eine brauchbare philosophische Theorie den Satz K beinhalten muss, der besagt, wie bestimmte Ausdrücke korrekt zu verwenden sind.

Sicherlich lässt sich eine Sprache denken, in der K nicht gilt, aber in dieser Sprache würden Ausdrücke wie beispielsweise ›Bild‹ und ›Darstellung‹ andere Bedeutungen besitzen beziehungsweise in anderen Folgerungszusammenhängen verwendet werden wie in der Umgangssprache.

Fassen wir kurz zusammen: Sätze der Art von K sind Sätze einer Theorie, die uns Auskunft über die Struktur unserer Rede über Bilder geben soll. Diese Theorie ist dabei empirisch neutral, insofern sie keine empirische Antwort darauf gibt, warum beispielsweise K gültig ist. Empirische Antworten sind dagegen solche, die behaupten, dass K aufgrund wahrgenommener Ähnlichkeiten oder bestimmter sozialer Strukturen gültig ist. Die Kompetenzbedingung im Sinne Lopes' ist dabei eine metatheoretische Bedingung, die von allgemeinen Bildtheorien fordert, dass sie mit K vereinbar sind und zu-

dem Aufschluss darüber geben sollen, warum K gilt.<sup>5</sup> Dieser Bedingung genügt beispielsweise Goodmans symboltheoretischer Ansatz nicht. Gleiches gilt beispielsweise auch für wahrnehmungspsychologische Theorieansätze, in denen man sich allein mit der Augenbewegung beim Bildbetrachten auseinandersetzt und in denen keine weiteren Schlüsse betreffs K gezogen werden. Theorieansätze der genannten Art sind somit hinsichtlich K neutral. Ob sie hinsichtlich der Gültigkeit Ks erweitert werden können, ist eine interessante Frage, die im Einzelfall nicht leicht zu beantworten ist. Der wichtigere Aspekt der Kompetenzbedingung ist für Lopes gleichwohl der, dass eine brauchbare allgemeine (!) Bildtheorie der Kompetenzbedingung gerecht werden muss und dabei hat er vor allen Dingen rivalisierende philosophische Bildtheorien im Blick. Nicht weniger wichtig, aber bisher noch nicht beachtet ist überdies der Aspekt, dass mit K zudem Bilder von sprachlichen Ausdrücken unterschieden werden sollen.<sup>6</sup> Das führt zu folgender Frage: Soll die Kompetenzbedingung nur auf Bildtheorien oder auch auf Sprachtheorien angewendet werden?

Will man diese Frage beantworten, sollte zuerst einmal klar sein, dass man sich hier auf einer Metaebene zweiter Stufe befindet, da man Auskunft über Eigenschaften von Theorien geben muss. Die Frage lässt sich aber in gewisser Hinsicht in der Objektsprache ausdrücken, wenn stattdessen gefragt wird: Wie unterscheidet sich das Verstehen sprachlicher Ausdrücke von dem Verstehen von Bildern?

Wie kann ein Wittgensteinianer diese Frage beantworten? Der naheliegende Weg, um zu untersuchen, was ›X versteht eine Sprache‹ und was ›X versteht ein Bildsystem‹ heißt, ist ihm aufgrund des nicht umgangssprachlichen Ausdrucks ›Bildsystem‹ verbaut, der in den Augen des Wittgensteinianers ja nur ein Pseudoausdruck ist.

Der Wittgensteinianer wird uns Geschichten erzählen und Beispiele geben, die beschreiben, was es heißt, aufgrund eines Bildes Personen oder Gegenstände zu erkennen. Aber all dies wird ihn nicht zu den Thesen führen: K gilt oder K gilt nicht für sprachliche Ausdrücke. Der bereits mehrfach genannte Grund ist, dass der Ausdruck ›Bildsystem‹ keinen Platz in seiner ›Theorie‹ hat. Ungeachtet dessen glaube ich, dass die Wittgensteinsche Therapie hilfreich zur Beantwortung der Frage ist: Ist eine Erweiterung der Kompetenzbedingung auch auf Sprachtheorien möglich oder kann sie nur auf Bildtheorien angewendet werden?

Warum? Weil dadurch zuerst einmal deutlich macht, dass gar nicht klar ist, was verglichen werden soll, wenn man versucht, die Kompetenzbedingung auf Sprachtheorien zu erweitern. Sind es Wörter, Kennzeichnungen, Beschreibungen? Klar ist, dass wir einem Wort nicht auf dieselbe Weise an-

---

<sup>5</sup> Goodman gibt zwar Bedingungen dafür an, wie sich sprachliche von bildnerischen Zeichen unterscheiden, aber auf Fragen zur Generativität und Transferenz geht er nicht ein.

<sup>6</sup> Soweit ich die Sachlage überblicke, wird in allen breiter angelegten philosophischen Bildtheorien darauf eingegangen, wie sich sprachliche Zeichen von Bildern unterscheiden.

sehen, was es bedeutet, wie einem Bild. Wie sieht es aber mit Beschreibungen folgender Art aus?

›Die Farbe Mauve liegt zwischen Veilchenblau und Burgunderrot.

Personen, die mit Farben und üblichen Farbsystemen vertraut sind, werden aufgrund dieser Beschreibung wissen, wie die Farbe Mauve aussieht, auch wenn sie den Ausdruck bis dato nicht kannten. Dagegen könnte man einwenden, dass man eine Farbtafel der Farbe Mauve neben einen Stoff legen und damit sehen kann, ob der Stoff die Farbe Mauve besitzt, die Beschreibung dagegen nicht, und gilt Gleiches nicht auch für Bilder? Sollen wir diesen Vergleich ernst nehmen? Ich sehe, was ein Bild darstellt, aber ich muss einen Text lesen, um zu wissen, was er bedeutet. Heißt das nun nicht nur, dass ich im Vergleich zum Verstehen eines Textes häufig schneller erfasse, was ein Bild zeigt?

An dieser Stelle ist es hilfreich, auf eine prinzipielle Eigenart von Bildern aufmerksam zu machen:

(S) Bilder werden mit den Augen wahrgenommen und stellen Ansichten von Dingen dar, die wir prinzipiell mit den Augen wahrnehmen können.<sup>7</sup>

Gleiches gilt für Texte nicht. Aber daraus folgt nicht, dass Generativität und Transferenz nicht auch auf Sprachen übertragbar sind beziehungsweise die Kompetenzbedingung nicht auch für Sprachtheorien gilt.

Spätestens hier sind wir wieder bei der Debatte zwischen Ähnlichkeits- und Zeichentheoretikern angekommen. Grund hierfür ist, dass der Ähnlichkeitstheoretiker ja sagen will, dass im Gegensatz zu Zeichentheorien seine Theorie der Kompetenzbedingung gerecht wird. Denn nur er, so der Ähnlichkeitstheoretiker, kann einfach und überzeugend erklären, wie es aufgrund der Ähnlichkeit zur Generativität und Transferenz bei Bildern kommt und warum Generativität und Transferenz bei Texten nicht vorliegt.

Nehmen wir an, der Ähnlichkeitstheoretiker hat recht, hat er damit eine noch nicht explizierte Sprachregel entdeckt im Sinne von:

Wenn eine Person mit einer Sprache vertraut ist, folgt daraus nicht notwendigerweise, dass sie alle Ausdrücke, die ihr bekannte Gegenstände etc. bezeichnen, sogleich versteht.<sup>8</sup>

Wohl nicht. Die Gültigkeit dieser Regel ist unabhängig von der These des Ähnlichkeitstheoretikers. Aus dieser Regel folgt aber auch nicht die Unhaltbarkeit einer konventionalistischen Auffassung. Denn genauso gilt:

---

<sup>7</sup> Diese These wird von vielen Bildtheoretikern (z.B. KULVICKI 2006; SCHOLZ 2004; WALTON 1990) kritisiert, aber für den Kernbereich der Bilder, um den es mir hier geht, trifft sie m.E. zu.

<sup>8</sup> Diese These widerspricht nicht K, da K die Abschwächung ›im Allgemeinen‹ enthält (vgl. S. 6).



Wenn eine Person mit einem Bildsystem vertraut ist, folgt daraus nicht notwendigerweise, dass sie alle Bilder, die ihr bekannte Gegenstände etc. darstellen, sogleich versteht.

Wie sieht es aber mit folgender These aus:

Eine Person ist mit einer Sprache genau dann vertraut, wenn sie im Allgemeinen neue Ausdrücke der Sprache sogleich verstehen kann, insofern ihr die bezeichneten Gegenstände etc. bekannt sind.

Diese These ist falsch, weil man einem neuen Wort aus einer Sprache, die man beherrscht, nicht sogleich ansieht, was es bezeichnet. Wie aber schon bemerkt, legt einen nichts darauf fest, Wörter als Vergleichsobjekte zu Bildern zu wählen. Vielmehr gilt im Allgemeinen die Regel:

(SK) Eine Person P ist mit einer Beschreibungssprache S\* genau dann vertraut, wenn die Person P im Allgemeinen (SG) sogleich eine neue Beschreibung aus S\* verstehen kann, insofern ihr die dargestellten Gegenstände etc. bekannt sind und (ST) sie ihr unbekannte Gegenstände etc. aufgrund einer Beschreibung aus S\* sogleich erkennen kann.

Die Einschränkung auf eine Beschreibungssprache wähle ich deshalb, weil ja auch K nicht auf alle Bilder anwendbar sein soll, sondern nur auf darstellende.<sup>9</sup> Ist aber SK intuitiv überzeugend? In einer starken Lesart von 'beherrscht eine Beschreibungssprache' trifft dies sicherlich zu. Denn wenn man davon ausgeht, dass eine Person nur dann eine Sprache beherrscht, wenn ihr das komplette Vokabular und alle grammatikalischen Regeln (Erzeugungsregeln von beliebig korrekten komplexen Ausdrücken) der Sprache bekannt sind beziehungsweise sie letztere korrekt anwenden kann, folgt, dass die Person jede neue Beschreibung sowohl im Sinne von SG und ST versteht. Denn aus dem Beherrschen eines Ausdrucks folgt nicht nur, dass man weiß, auf welche Ereignisse etc. er bisher angewendet wurde, sondern auch auf welche noch unbekanntes künftigen Ereignisse.<sup>10</sup> Offensichtlich aber widerspricht eine solche starke Lesart der Intuition der meisten kompetenten Sprecher. Wörterbuch und Grammatik einer Sprache sind häufig im Bücherregal von Personen anzutreffen, die wir gerade als besonders kompetente Sprecher ansehen. Das

---

<sup>9</sup> Ein Grund, warum sich bei Lopes eine Überlegung in diese Richtung nicht findet, könnte sein, dass seiner Auffassung nach eine Theorie, die Bilder analog zu sprachlichen Beschreibungen auffasst, zulassen muss, dass ein Portrait (z.B. eines eineiigen Zwillinges) genauso wenig den Bezug genau auf einen Zwilling festlegt wie eine (äußerliche) Beschreibung eines der Zwillinge (LOPES 1996: 5.1). Dies ist für Lopes aber unakzeptabel, daher vertritt er eine kausale Theorie im Sinne Kripkes.

<sup>10</sup> Erinnerung sei in diesem Zusammenhang auch daran, dass das sogenannte Regelfolgenparadox nur unter der Lesart Kripkes ein Paradox für Wittgenstein darstellt (vgl. BAKER/HACKER 1984).