

Bildverstehen und Sprache

Steffen-Peter Ballstaedt

1. Das Problem

Was läuft im Kopf ab, wenn ein Betrachter eine Abbildung, z.B. ein Foto oder eine repräsentationale Zeichnung versteht? Und welche Rolle spielt die Sprache beim Bildverstehen? Letztere Frage hat in der kognitiven Psychologie immer wieder zu Kontroversen geführt. Als Beispiel kann die Auseinandersetzung um die duale Kodierung dienen, in der Bilder grundsätzlich imaginal *und* verbal verarbeitet werden (Paivio 1971). Auch in der Kunstwissenschaft wird dieses Thema unter der Überschrift "Die Sprachlichkeit der Bilder" aufgegriffen (We-weder 1985).

2. Bildverstehen als Forschungsthema

Das Textverstehen wurde jahrzehntelang in der kognitiven Psychologie ausgiebig beforscht, während das Bildverstehen ein eher randständiges Thema blieb. Nun mag man sich fragen, ob das Verstehen von Bildern überhaupt neue Forschungsbemühungen notwendig macht. Reicht nicht für die "höhere" *Verarbeitung* das umfangreiche Wissen über Textverstehen aus? Viele Erkenntnisse z.B. über die Wirkung von Schemata oder über kohärenzbildende Prozesse sind sicher übertragbar. Und reicht für die "niedere" *Verarbeitung* nicht die Wahrnehmungspsychologie aus? Die Prozesse bei der visuellen Wahrnehmung dürften sich kaum dramatisch von denen der Bildwahrnehmung unterscheiden. Aus der Neuropsychologie wissen wir, daß das natürliche Sehen und die Bildbetrachtung in denselben neuronalen Arealen stattfinden. Es gibt keine eigenen Module für die Bildverarbeitung, während die Verarbeitung der gesprochenen Sprache und der Schrift in eigenen Modulen abläuft. Ist das Bildverstehen nur ein Sonderfall des Sehens mit reduzierten visuellen Vorgaben?

Vor dieser schlichten Sicht behütet uns die Semiotik, die darauf aufmerksam macht, daß Bilder nicht einfach einen besonderen Typ von *visuellen Stimuli* darstellen, sondern *ikonische Zeichen* sind, die ein Sender benutzt, um einem Empfänger etwas zu zeigen, um etwas auszudrücken und etwas zu bewirken. Greifen wir auf das semiotische Organonmodell von Bühler (1934) zurück, das Darstellungs-, Ausdrucks- und Appellfunktion bei Zeichen unterscheidet, so fällt auf, daß die *Darstellungsfunktion* in der natürlichen Einstellung gegenüber Abbildungen dominiert, sie gelten oft als Realitätersatz. Die anderen kommunikativen Funktionen – Ausdruck und Appell – werden gern übersehen. Das Verstehen eines Bildes umfaßt aber alle drei Funktionen. Weidenmann (1988) unterscheidet

das *ökologische Verstehen* der Inhalte vom *indikatorischen Verstehen* der beiden anderen Zeichenfunktionen.

So führen Bilder ein merkwürdiges Doppelleben als Surrogat und Kommunikat. Bei einem Text handelt es sich zweifellos um ein Zeichensystem aus Buchstaben, Morphemen oder Wörtern. Eine Verwechslung mit der Realität wäre ein psychopathologischer Fall (wie bei dem Schizophrenen, der auf ein Papier "Schnitzel" schreibt und es dann genüßlich verspeist). Bei Abbildungen wird demgegenüber der Zeichencharakter gern vergessen.

3. Stufen der Bildverarbeitung

In der kognitiven Psychologie ist die Einteilung in Stufen der Verarbeitung beliebt, auch wenn es sich meist um ein Kontinuum von Prozessen handelt. Aber zur Orientierung sind die Stufen recht nützlich (Pettersson 1988; Weidenmann 1993; Ballstaedt, im Druck).

Präattentive Verarbeitung

Ein Blick (bereits 100 msec) auf ein Bild genügt, um ein Objekt oder eine Szene wahrzunehmen, zu erkennen bzw. zu kategorisieren (Biedermann 1981). Bewirkt wird dies durch angeborene Verarbeitungsprozeduren, die automatisch und parallel eine perzeptuelle Organisation schaffen. Beteiligt sind die sogenannten *Gestaltgesetze* und dann die *Mustererkennung* aufgrund visueller Schemata (anschauliche Begriffe). Die Funktion dieser Stufe besteht in einer schnellen räumlichen Orientierung, deren biologischer Nutzen auf der Hand liegt. Die Verarbeitung auf dieser Stufe ist weitgehend unbewußt und willentlich nicht beeinflussbar. Daß es diese Stufe der Verarbeitung gibt, wird erst bewußt, wenn keine prägnante und eindeutige perzeptuelle Organisation zustande kommt. Beim natürlichen Sehen ist dies selten der Fall, es kann aber durch besondere artifizielle Vorgaben für das visuelle System provoziert werden (z.B. Kippfiguren).

Ein Bildtyp, dessen Verstehen auf dieser Stufe der Verarbeitung bleibt, sind die *Piktogramme*. Sie sollen auf einen Blick ein Konzept aktivieren oder eine Handlung auslösen, ohne Auswertung durch Augenbewegungen und andere Verarbeitungsprozesse. Ob dieses Ziel immer erreicht wird, sei hier dahingestellt.

Attentive Verarbeitung

Auf der zweiten Stufe beginnt eine *Detailauswertung*, deren beobachtbarer und meßbarer Indikator die *sakkadischen Augenbewegungen* sind (Issing et al. 1985). Sie sind seriell: Ein Bildpunkt nach dem anderen wird fixiert und damit in die Stelle des schärfsten Sehens auf der Retina gebracht. Die Sakkaden und Fixa-

tionen lassen sich registrieren und das Muster gibt Auskunft über die Verteilung der Aufmerksamkeit. Die Augenbewegungen sind zu einem kleinen Teil reflektorisch gesteuert, sie reagieren z.B. auf unerwartete oder hervorgehobene Bildetails. Auf diese Weise läßt sich die Auswertung auf dieser Stufe über die Bildgestaltung (*cueing*) in Grenzen beeinflussen. Der größere Anteil ist jedoch bewußt kontrollierbar und durch Interessen, Vorwissen oder Aufgaben gesteuert. Der Betrachter entscheidet, was er genauer, d.h. öfter und länger anschaut. Die Funktion dieser Verarbeitungsstufe liegt in der *aktiven Exploration* der visuellen Umgebung, hierzu gibt es domainenspezifische visuelle Routinen.

Bilder, die auf dieser Stufe verstanden werden müssen, sind z.B. *mikroskopischen Aufnahmen* oder *Computertomografien*. Ihre medizinische bzw. neurologische Auswertung erfordert ein genaues Scanning unter Steuerung des Vorwissens, um bekannte Strukturen zu entdecken und Abweichungen zu erkennen.

Elaborative Verarbeitung

Diese dritte Ebene ist schwieriger zu fassen, sie kann auch als Vertiefung der zweiten Ebene angesehen werden. Hier geht es um durch das Bild ausgelöste, aber über es hinausgehende Schlußfolgerungen, Assoziationen, Vorstellungen, die durch vorhandenes Wissen (Schemata) ausgelöst werden. Dies ist die Stufe, auf der ein Betrachter sich die Frage stellt: Mit welcher Absicht wird mir dieses Bild gezeigt? Hier vor allem findet das indikatorische Verstehen statt. Funktion dieser Verarbeitungsstufe ist die multiple Integration in vorhandene Wissensbestände.

Als Beispiel für elaborative Verarbeitung können künstlerische Bilder dienen, z.B. Dürers bekannter Kupferstich *Melencolia*. Über dieses Bild gibt es viele Abhandlungen und Bücher, da es vieldeutige Bestandteile enthält. Die elaborative Stufe umfaßt unter Umständen eine jahrelange Auseinandersetzung mit einem Bild. Im Prinzip gilt für diese Stufe "open end". Untersuchungen belegen, daß diese Stufe der Bildverarbeitung und des -verstehens nicht selbstverständlich ist, sondern gelernt und trainiert werden muß (Schmidt/McLaughlin/Leighen 1989). Sie setzt vor allem voraus, daß man sich von dem Erkennen von Inhalten löst und auch auf die formale Gestaltung achtet.

Diese drei Stufen finden sich in allen Ansätzen zur Bildverarbeitung. Dabei ist nicht zu übersehen, daß es sich um ein Kontinuum zunehmenden Verarbeitungsaufwands und zunehmender *Verarbeitungstiefe* handelt: Man kann nur einen Blick auf ein Bild werfen (präattentive Verarbeitung), man kann ein Bild ausführlich explorieren (attentive Verarbeitung) und man kann schließlich weiterführende Gedanken investieren (elaborative Verarbeitung). Von Stufe zu Stufe nimmt der Anteil der top-down-Verarbeitung zu.

4. Zur Rolle der Sprache beim Bildverstehen

Kommen wir nun auf unsere zentrale Frage zurück: Welche Rolle spielt die sprachliche Verarbeitung in diesem Kontinuum des Bildverstehens. Dazu einige Thesen, die mit empirischen Untersuchungen untermauert werden können, die wir hier nicht ausführlich referieren (ausführlich Engelkamp 1990).

1. Das Erkennen von Objekten und Szenen ist zunächst schemagesteuert und vorsprachlich. Einfache Bilder werden nicht automatisch benannt, komplexe Szenen nicht spontan beschrieben.

Durch Introspektion ist schwer herauszubekommen, ob bei der Betrachtung von Bildern sprachliche Prozesse notwendig beteiligt sind: Manchmal kommen einem Worte oder Sätze in den Kopf, manchmal offensichtlich nicht. Um die formulierte These zu überprüfen, muß man sich eine pfiffige experimentelle Anordnung ausdenken. Eine Gruppe von Forschern um Nelson hat dies bereits in den siebziger Jahren getan (Nelson/Brooks 1973; Nelson/Reed 1976). Sie benutzen die in der Lernpsychologie gebräuchliche Methode des *Paarlernens*: Eine Liste von den Itempaaren wird nacheinander dargeboten, meist Wort-Wort, beim zweiten Durchgang bekommen die Versuchspersonen nur das erste Reizwort und müssen das zweite Reaktionswort antizipieren, erst dann bekommen sie zur Kontrolle das Paar geboten. Die Liste wird so lange dargeboten, bis alle Elemente viermal korrekt antizipiert sind. Als Lernleistung wird die durchschnittliche Anzahl der Fehler bis zu diesem Kriterium gemessen.

Die Experimentatoren untersuchten zwei unabhängige Variablen in ihrer Wirkung auf die *Lernleistung* (als abhängige Variable):

Lautähnlichkeit: Die Wörter einer Liste waren zwar semantisch unverbunden, aber lautähnlich (sie hatten die gleichen An- und die gleichen Auslaute). Die Wörter in der anderen Liste waren lautlich völlig verschieden.

Darbietungsform: Es wurden Wort-Wort-Paare und Bild-Wort-Paare verwendet. Bei den Bildern wurden noch einmal zwei Versuchsgruppen unterschieden: Die eine mußte nach dem Reizbild das assoziierte Wort antizipieren. Bei der anderen Gruppe wurde das Bild vom Versuchsleiter benannt und der Name mußte ausgesprochen werden. Wir haben damit drei Bedingungen, unter denen die Paare gelernt werden müssen: Wort-Wort, benanntes Bild-Wort, Bild-Wort.

Tabelle 1: Mittlere Fehlerzahl als Funktion der Lautähnlichkeit der Reizwörter und der Darbietungsart (nach Nelson/Brooks 1973).

Lautähnlichkeit	Darbietungsart		
	Bild	benanntes Bild	Wort
groß	17,81	40,19	69,75
klein	21,44	33,31	39,56

Die Befunde zeigen einen deutlichen Effekt der beiden unabhängigen Variablen. Eine große Lautähnlichkeit führt bei der Wort-Wort-Bedingung zu einer markanten Verschlechterung beim Lernen gegenüber geringer Lautähnlichkeit. Dieser Effekt kommt durch Interferenzen zwischen den Reizwörtern zustande. Er ist schwächer ausgeprägt in der Gruppe mit den benannten Bildern, verschwindet aber in der Bild-Bedingung (hier ist der Effekt sogar umgekehrt). Dies zeigt, daß Bilder nicht implizit benannt werden, sonst müßte sich auch in der Bild-Darbietung ein beeinträchtigender Interferenzeffekt zeigen. Wer ein Bild betrachtet – und hier geht es um einfache Bilder von Objekten –, der aktiviert nicht spontan die Wortform. Es geht also ein direkter Weg vom Bild zum Konzept. Visuelle Informationen konstituieren Konzepte, dagegen sind Wörter nur Stellvertreter für Konzepte. Es gibt weitere empirische Untersuchungen mit anderen Methoden, die diesen Befund bestätigen.

2. Die Versprachlichung von Bildern ist aufgabenabhängig. Entweder wird sie durch Instruktion zur Benennung, Beschreibung oder Interpretation hervorgerufen. Oder sie stellt eine bewußte Strategie zur vertiefenden Verarbeitung dar.

Diese These wird nicht mit einem Experiment belegt, da sie wohl niemanden sonderlich überrascht. An einer teilweisen Verbalisierung von Bildern kommt ein Betrachter nicht vorbei, wenn er über ein Bild nachdenken und vor allem darüber kommunizieren möchte. Die höheren Verarbeitungsprozesse sind weitgehend an Sprache gebunden, was nicht ausschließt, daß es auch ein anschauliches Denken gibt (Arnheim 1972; Babbitt 1982).

3. Wenn Bilder benannt oder beschrieben werden, dann verbessert sich die Behaltensleistung. Verbalisierung ist eine effektive Strategie der Bildverarbeitung und des Wissenserwerbs mit Bildern.

In psycholinguistischen Untersuchungen konnte man zeigen, daß sich die Rekognitionsleistung signifikant verbessert, wenn eine Aufgabe die Verbalisierung komplexer Bilder verlangt (Bartlett/Till/Levy 1980). Umgekehrt konnte man nachweisen, daß die Unterdrückung der Verbalisierung durch ablenkende Aufgaben – z.B. während der Bildbetrachtung in Dreierschritten rückwärts zu zählen – die Rekognitionsleistung beeinträchtigt (Loftus 1972; Loftus/Kallman 1979).

Eine zweite Gruppe von empirischen Belegen liefern Trainingsprogramme zum Umgang mit Visualisierungen. Sie enthalten als eine effektive Strategie, die wichtigsten Aussagen des Bildes mündlich oder schriftlich zusammenzufassen (Moore 1993). Besonders gilt dies für das indikatorische Verstehen der kommunikativen Botschaft von Bildern. Die intensive Auseinandersetzung und tiefere Verarbeitung von Bildern ist ein Skill, der z.B. Studenten und Studentinnen der Kunstgeschichte beigebracht werden muß.

4. Es besteht ein Zusammenhang zwischen Augenbewegungen, Verbalisierung und Behalten. Die Verteilung der visuellen Aufmerksamkeit durch Fixationen von Bilddetails schlägt sich in einem selektiven Behalten nieder.

Entscheidend für die Behaltensleistung ist die Anzahl der Fixationen (Loftus 1972). Je wichtiger und je komplexer ein Bilddetail, desto öfter wird es fixiert.

Während einer Fixation wird nicht nur die visuelle Information analysiert, sondern es laufen Prozesse der Kategorisierung und der Verbalisierung ab. Aus dem Fixationsmuster lassen sich die Inhalte einer freien Reproduktion vorhersagen.

5. *Durch den sprachlichen Kontext (Legende, Kommentar, Begleittext usw.) läßt sich die Auswertung von Bildern in begrenztem Ausmaß steuern. Dies ist wichtig für die Gestaltung instruktionaler und persuasiver Text-Bild-Beziehungen.*

Ein komplexes Bild läßt der Auswertung etliche Freiräume: Die Blickbewegungsmuster verschiedener Betrachter sind individuell unterschiedlich und ebenso die Kategorisierungen und Beschreibungen. Ein Bild ist mehrdeutig, offen für unterschiedliche Sehweisen und Interpretationen. Eine Reihe von Untersuchungen hat gezeigt, daß ein *sprachlicher Kontext* die Auswertung eines Bildes beeinflussen kann. Wir unterscheiden drei Möglichkeiten (nach Zimmer 1983):

- ein zusätzlicher Kommentar kann die Einstellung des Betrachters zur Abbildung verändern (z.B. ein Bild als Urlaubsfoto oder als Werbefoto),
- eine sprachliche Anleitung lenkt die Augen auf bestimmte Bilddetails (Akzentuierung) oder legt bestimmte Blickpfade nahe,
- durch sprachliche Bezeichnungen kann die Kategorisierung der Perzepte vorgegeben werden: Sieht man eine Blume, eine weiße Rose, eine Kaiserin Augusta Viktoria? Je konkreter ein Objekt benannt wird, desto mehr visuelle Details werden behalten (Jörg 1978).

Das Bildverstehen kann sprachlich angeleitet werden, indem *vorher* eine Wahrnehmungseinstellung erzeugt wird, indem *während* der Betrachtung die visuelle Aufmerksamkeit gelenkt wird oder indem *nachher* die elaborative Verarbeitung beeinflusst wird (Wiseman/MacLeod/Lootsteen 1985). Damit wird die Kreativität des Sehens kanalisiert. Aber die Sprache kann ein Bild nie ganz vereinnahmen, ein Bild bleibt zu einem Teil im Wortsinn "unsagbar" oder "unbeschreiblich".

6. *Bilder sind multipel (multimodal) mental repräsentiert: Als Bildmarken in einem visuellen Langzeitgedächtnis, als teilweise sprachliche Kopie und als rein konzeptuelle Repräsentation. Welche Repräsentation genutzt wird, hängt von den Anforderungen der Aufgabe ab.*

Das Gedächtnis wird heute als Beiprodukt der Verarbeitung und nicht als eigene Stufe der "Ablagerung" von Endprodukten gesehen. Das bedeutet, daß Spuren von Verarbeitungsprozessen auf verschiedenen Stufen repräsentiert sind, je nachdem was der Betrachter mit dem Bild angefangen hat. Die verschiedenen Repräsentationen werden je nach Aufgabenstellung genutzt. Ob in diesem Zusammenhang der beliebte Begriff des *mental Modells* einen Fortschritt bringt, soll hier nicht erörtert werden.

Zusammenfassung

Auf den ersten Blick verarbeitet ein Betrachter eine Abbildung visuell und zumindest teilweise konzeptuell. Er versteht es, bevor eine sprachliche Verarbeitung einsetzt. Diese dient dann der kontrollierten tieferen Verarbeitung und natürlich der Kommunikation über ein Bild.

Bibliographie

- Arnheim, R.: *Anschauliches Denken. Zur Einheit von Bild und Begriff*. Köln: DuMont, 1972.
- Babbitt, B.C.: "Effect of task demands on dual coding of pictorial stimuli." In: *Journal of Experimental Psychology: Human Learning and Memory* 8 (1982), p. 73-80.
- Ballstaedt, St.-P.: *Bilder funktional gestalten. Seminarunterlagen*. Tübingen: DIFF, 1994.
- Ballstaedt, St.-P.: "Bildverstehen, Bildverständlichkeit. Ein Forschungsüberblick unter Anwendungsperspektive." In: Krings, H.P. (ed.): *Wissenschaftliche Grundlagen des technischen Schreibens* (im Druck).
- Bartlett, J./ Till, R.E./ Levy, J.C.: "Retrieval characteristics of complex pictures: Effects of verbal encoding." In: *Journal of Verbal Learning and Verbal Behavior*, 19 (1980), p. 430-449.
- Biedermann, I.: "On the semantics of a glance at a scene." In: M. Kubovy/ J. R. Pomerantz (eds.): *Perceptual organization*. Hillsdale, NJ: Erlbaum, 1981, p. 213-253.
- Bühler, K.: *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer, 1934.
- Engelkamp, J.: *Das menschliche Gedächtnis. Das Erinnern von Sprache, Bildern und Handlungen*. Göttingen: Hogrefe, 1990.
- Issing, L.J./ Beyland, U./ Haack, J.H./ Mickasch, H.D. (eds.): *Blickbewegungsforschung und Bildverarbeitung*. Berlin: Arbeitsgruppe Medienpsychologie/Medienpädagogik, 1985.
- Jörg, S.: *Der Einfluß sprachlicher Bezeichnungen auf das Wiedererkennen von Bildern*. Bern: Huber, 1978.
- Loftus, G.R.: "Eye fixations and recognition memory for pictures." In: *Cognitive Psychology*, 3 (1972), p. 525-551.
- Loftus, G.R./ Kallman, H.J.: "Encoding and use of detail information in picture recognition." In: *Journal of Experimental Psychology: Human Learning and Memory*, 5 (1979), p. 197-211.
- Moore, Ph.J.: "Metacognitive processing of diagrams, maps and graphs." In: *Learning and Instruction*, 3 (1993), p. 215-226.
- Nelson, D.L./ Brooks, D.H.: "Independence of phonetic and imaginal features". In: *Journal of Experimental Psychology*, 97 (1973), p. 1-7.
- Nelson, D.L./ Reed, V.S.: "On the nature of pictorial encoding: A levels-of-processing analysis." In: *Journal of Experimental Psychology: Human Learning and Memory*, 2 (1976), p. 49-57.
- Schmidt, J.A./ McLaughlin, J.P./ Leighton, P.: "Novice strategies for understanding paintings." In: *Applied Cognitive Psychology* 3 (1989), p. 65-72.
- Paivio, A.: *Imagery and verbal processes*. New York: Holt, Rinehart & Winston, 1971.
- Pettersson, R.: "Interpretation of image content." In: *Educational Communication and Technology Journal*, 36, (1988), p. 45-55.
- Weidenmann, B.: *Informierende Bilder*. München: Universität der Bundeswehr, 1993. [Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie, Nr. 29.]
- Weweder, R.: *Zur Sprachlichkeit von Bildern. Ein Beitrag zur Analogie von Sprache und Kunst*. Köln: DuMont, 1985.
- Wiseman, S./ MacLeod, C.M./ Lootsteen, P.J.: "Picture recognition improves with subsequent verbal information." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (1985), p. 588-595.