

Stephan Butzke:

***Ein schwieriges Verhältnis aus der Sicht der Semiotik:
Text und Bild***

Wird das Verhältnis von Äußerungen, „die sich eines natürlichsprachigen Zeichensystems“ bedienen, und denen, „die sich solcher non-verbaler Zeichensysteme“ bedienen, „deren Zeichen einerseits visuell wahrnehmbar sind und nicht nur als abbildendes Notationssystem [...] eines andersartigen Zeichensystems fungieren“¹, untersucht und dabei von einem *Text-Bild-Verhältnis* gesprochen, ist eine erste Annahme bereits einbegriffen: *Bilder sind keine Texte*. Die Üblichkeit des Redens von Text-Bild-Verhältnissen ist indes bezeichnend, da es in der Semiotik sonst häufig die Tendenz gibt, „alle kulturellen Systeme [...] als Sprachen“², und damit deren Äußerungen (Parole) als sprachliche Texte aufzufassen. Die genannte Annahme ist daher sicher oft nur unbeabsichtigt impliziert, was auf die Schwierigkeit der Diskussion deutet. Hier soll die Begriffsopposition Bild – Text *bewusst* beibehalten werden.

So besteht zunächst das Problem, die Fähigkeit, außersemiotische Wirkungen hervorzurufen, bei Bildern und Texten zu unterscheiden. Bilder können schon auf der Ebene des Signifikanten, unabhängig von möglichen Bedeutungen, z.B. affektive oder neurophysiologische Reaktionen auslösen, während ein sprachlicher Text eine solche Leistung nur über den Umweg seiner Signifikate erbringen kann und dabei darauf angewiesen ist, beim Empfänger entsprechende *Bilder* zu evozieren. Das Bild ist also in der Lage, Informationen im Sinne eines Signals, das gleichzeitig als Zeichen fungieren kann, aber nicht muss, einem menschlichen Empfänger direkt zu übermitteln.

Weniger eindeutig gestaltet sich der Unterschied bei der Fähigkeit beider Systeme, ästhetische Wirkungen hervorzubringen, da dies beiden beim Ausdruck sowie beim Inhalt möglich ist: ein Bild kann in der farblichen Gestaltung, der syntaktischen Verknüpfung seiner Elemente und in seiner Bedeutungsstruktur als ästhetisch ansprechend empfunden werden, so, wie die Sprache mit Hilfe stilistischer Mittel ihre Signifikanten reizvoll gestalten kann und es bezüglich der Ästhetik der Bedeutungen wahrscheinlich einfacher hat als das Bild.

1 Michael Titzmann: *Theoretisch-methodologische Probleme einer Semiotik der Text-Bild-Relationen*. In: *Text und Bild, Bild und Text..* DFG-Symposium 1988. Hrsg. von Wolfgang Harms. Stuttgart 1990 (= Germanistische Symposien. Berichtsbände XI), S. 368-384, Zitat S. 368.

2 Erika Fischer-Lichte: *Ästhetische Erfahrung. Das Semiotische und das Performative*. Tübingen, Basel 2001. Zitat S. 9. Scholz spricht in seinen Studien zur Emblemik vom Emblem als einem *Wort-Bild-Text*. Bernhard F. Scholz: *Emblem und Emblempoetik. Historische und systematische Studien*. Berlin 2002 (Allgemeine Literaturwissenschaft – Wuppertaler Schriften Bd. 3).

Ein dritter außersemiotischer Sachverhalt, die mnemonische Qualität, ist wieder einfacher handhabbar.³ Es lässt sich kaum bestreiten, dass Bilder und folglich die einem Bild möglicherweise beigegebenen Bedeutungen einprägsamer sind als Worte.

Die eigentliche Diskussion um die *Relation* von Texten und Bildern fängt natürlich erst dann an, wenn man die Semiosefähigkeit beider Systeme in Erwägung zieht und bedenkt, dass jedes Signifikat eines nichtsprachlichen Zeichens nur dann ein solches ist, wenn es durch einen sprachlichen Signifikanten bedeutet werden kann.⁴ Außerhalb sprachlicher Artikulierbarkeit als Voraussetzung für Bewusstseinsfähigkeit ist Bedeutung nicht denkbar.⁵ Diese Aussage kann durchaus als Tautologie verstanden werden, da es möglich ist, Denkbarkeit und Artikulierbarkeit synonym zu benutzen. In dem Moment, in dem ein Produzent oder ein Rezipient einem Bild oder einem Element desselben eine Bedeutung zuschreibt, hat er den bildlichen Ausdruck schon durch einen sprachlichen ersetzt. Demnach müsste, streng semiotisch gesehen, die Sprache als das überlegene Zeichensystem eingestuft werden.

Eine solche Überlegenheit könnte man genauso gut in dem bei sprachlichen Äußerungen geringeren Umfang des Anteils sehen, der keine Bedeutungen generiert.⁶ Bei erweitertem Blickwinkel ist hierin möglicherweise aber gerade ein Vorteil des Bildes und ein Grund dafür, dass Bilder überhaupt produziert werden, zu erkennen. Denn gerade die bedeutungslosen Anteile einer bildlichen Äußerung sind vermutlich besonders geeignet, die oben beschriebenen außersemiotischen Wirkungen hervorzurufen.

Wenn ein Bild jedoch mit der Intention, kommunikativ, d.h. bedeutungsübermittelnd, zu sein, produziert wird – und diese Funktion haben wohl die meisten Bilder zumindest teilweise – ergibt sich neben dem geringeren Anteil zeichenhafter Elemente die Schwierigkeit, diese als solche erkennbar zu machen. Denn während die Sprache aufgrund ihrer hohen Kodiertheit relativ mühelos Korrelationen zwischen der Ausdrucks- und Inhaltsebene herstellt und ein sprachliches Element deshalb sofort als Zeichen identifiziert wird, muss das Bild einen Kode erst instaurieren.⁷ Das ist in dreifacher Hinsicht problematisch: es ist nicht geregelt, welche Elemente bedeutungsdifferenzierend und welche bedeutungstragend sind, und noch weniger

³ Vgl. Wolfgang Neuber: *Locus, Lemma, Motto. Entwurf zu einer mnemonischen Emblematiktheorie*. In: *Ars memorativa. Zur kulturgeschichtlichen Bedeutung der Gedächtniskunst 1400-1750*. Hrsg. von Jörg Jochen Berns und Wolfgang Neuber. Tübingen 1993 (=Frühe Neuzeit Bd. 15), S. 351-372. Sowie: Ursula Kocher: *Der Dämon der hermetischen Semiose – Emblematik und Semiotik*. In: *Bildersprache verstehen*. Hrsg. von Ruben Zimmermann. München, Paderborn 2000, S. 151-167.

⁴ Vgl. Roland Barthes: *Elemente der Semiologie*. Frankfurt am Main 1983, S. 37.

⁵ Titzmann, S. 371.

⁶ Ebd. S. 375.

steht fest, *welche* Bedeutung ein mögliches Zeichen trägt.⁸

Ein erster Schritt einer solchen Kode-Instaurierung ist die Anwendung des Relationsprinzips ikonischer Zeichen: Objekte im Bild werden aufgrund ihrer Ähnlichkeit mit entsprechenden Signifikaten korreliert. Dazu muss sich das entsprechende Objekt aus im menschlichen Wahrnehmungsapparat gespeicherten strukturellen Mustern zusammensetzen lassen, also nicht schon unbedingt bekannt sein.⁹ Da das hier gespeicherte Wissen über die Welt recht individuell geprägt ist, wird der so in Gang gesetzte Interpretationsprozess zu einem recht individuellen Akt. Welches Bildelement z.B. zuerst als Signifikant wahrgenommen wird, variiert von Interpret zu Interpret. Genau dieses Element dient aber anschließend als Maßstab für die Deutung der übrigen Elemente.¹⁰ Ausgehend von einer erkannten Bedeutung interpretiert der Betrachter die verbleibenden potentiellen Signifikanten so, dass sich eine sinnvolle Bedeutungsstruktur ergibt. Während in der Sprache der Signifikant mit seinen materiellen Eigenschaften ein Zeichen markiert, wird ein Bildelement also erst aufgrund einer beim Interpretieren zugewiesenen Bedeutung zum Signifikanten (Kode-Instaurierung). Gleichfalls ist dieser Prozess abhängig von der individuellen Wahrnehmungs- und Wissensstruktur des Rezipienten (und Produzenten), so dass selbst bei Übereinstimmung des zuerst als Signifikant wahrgenommenen Objektes sowie des zugeordneten Signifikats das Endergebnis nie einheitlich sein wird.

Die so gegebene Ambiguität kann selbstverständlich sprachlichen Texten gleichermaßen zu Eigen sein. Vor allem die stilistischen Mittel ästhetischer Texte sind ähnlich schwach kodiert, ganz zu schweigen von eventuellen Bedeutungen evozierter Bilder, bei denen schon die (evozierte) Ausdrucksebene von Mensch zu Mensch variiert. Der entscheidende Unterschied zum Bild liegt jedoch darin, dass ein Text immer unmittelbare und explizite konventionell vorgegebene Bedeutungen trägt. Diese können die Interpretation nur sekundär im Text angelegter Bedeutungen unterstützen.

Daraus lässt sich nun leicht schlussfolgern, dass ein Text diese Unterstützung genauso einem Bild zur Verfügung stellen kann und sogar muss, oder, anders herum gesagt, ein Bild, das sekundäre Bedeutungen generieren soll, ist auf die Kooperation mit einem Text angewiesen. Dabei ist die direkte und visuell wahrnehmbare Nachbarschaft zu einem Text, wofür das frühneuzeitliche Emblem das wichtigste Beispiel wäre, nur ein Sonderfall. Liegt eine solche

⁷ Vgl. Umberto Eco: *Entwurf einer Theorie der Zeichen*. München 1991, S. 286 f.

⁸ Titzmann, S. 377.

⁹ Ebd. S. 378.

¹⁰ Ebd. S. 379.

direkte Nachbarschaft nicht vor, erfolgt die Interpretation des Bildes im Rahmen all derjenigen vom Betrachter im Laufe seines Lebens schon rezipierten und im Gedächtnis gespeicherten Texte, deren Bedeutungen sich sinnvoll in die potentielle Bedeutungsstruktur des Bildes einfügen lassen.¹¹

Wenn ein Bild z.B. eine generalisierende Aussage treffen soll, muss dem Betrachter das sprachlich mitgeteilt werden, da ein Bild primär immer nur spezifische Objekte oder Ereignisse darstellt.¹² Ebenso unterliegt das im Bild dargestellte einer synchronen Zustandhaftigkeit und kann erst durch einen sprachlichen Text als Teil einer diachronen Zustandsfolge eingeordnet werden. Schließlich muss wohl kaum erwähnt werden, dass Metaaussagen über sekundäre Bedeutungen im Bild nicht einmal mehr indirekt darstellbar sind und allein durch Texte kommuniziert werden müssen.

Konzentriert man sich nun aber unabhängig von möglichen sekundären Bedeutungen oder gar Metaaussagen auf die reine Darstellungsfunktion, könnte sich das Gewicht wieder zugunsten des Bildes verschieben und ein weiterer Grund dafür, dass Bilder überhaupt produziert werden, deutlich werden. Das im Bild visuell Wahrnehmbare ist nicht nur dargestellt, sondern vielmehr *abgebildet* und damit durch Konkretheit auf der einen und Vollständigkeit auf der anderen Seite gekennzeichnet. Im Text werden visuelle Elemente nur impliziert und müssen daher mit Abstraktheit auf der einen und Unvollständigkeit auf der anderen Seite arbeiten, wodurch es zu der oben erwähnten Verschiedenheit evozierter Bilder kommt, denn jeder Leser konkretisiert die Abstraktheit und ergänzt die Unvollständigkeit individuell. Daher ist es nicht verwunderlich, dass z.B. Buchillustrationen in verschiedenen Ausgaben desselben Textes nie identisch sind. Eine solche Visualisierung des Textes ist immer nur eine spezifizierende Konkretisation, die allerdings bei Bedarf den Bezug des Textes auf mögliche Referenten erleichtern kann. Umgekehrt lässt sich ein Bild sprachlich immer nur unvollständig beschreiben.¹³ Das deutet nun wieder auf den erwähnten Aspekt des großen Anteils nichtsignifikanter, also nicht artikulierbarer Elemente beim Bild.

Abschließend soll noch ein Blick auf den schon erwähnten Sonderfall der direkten Bild-Text-Koexistenz geworfen werden, da sich die jeweils speziellen Eigenschaften beider Systeme, wenn auch implizit, so doch noch einmal recht verständlich zeigen lassen. Dabei müssen zwei Arten einer eventuellen Koexistenz unterschieden werden: die Sprechinstanz kann bildlich dargestellt sein, so dass der Text im Bild eingebettet ist, oder die Sprechinstanz des Textes

11 Ebd. S. 375.

12 Ebd. S. 379.

13 Ebd. S. 380.

befindet sich außerhalb des Bildes. Bei dieser Variante muss man weiterhin die Möglichkeit einer Gleichrangigkeit von Bild und Text von der einer Einbettung des Bildes in den Text unterscheiden.¹⁴

Im ersten Fall werden die möglichen Textbedeutungen durch die Bedeutungen des Bildes begrenzt. Dadurch, dass das Bild den Sprecher und (eventuell) den Adressaten sowie Ort, Zeit und andere Umstände der Sprechsituation identifiziert, werden nur dem entsprechende Signifikate des sprachlichen Textes selektiert und eventuell sogar referenzialisiert. Dieses Verhältnis kehrt sich um, wenn sich die Sprechinstanz außerhalb des Bildes befindet. Hier werden Textbedeutungen zum strukturierenden Angelpunkt der dem Bild zugeschriebenen Signifikate.¹⁵ Trotz der enormen Vieldeutigkeit bildlicher Signifikanten ist dieser Semioseprozess aber nicht einer völlig unbegrenzten Willkürlichkeit unterworfen. Eine gewisse Kompatibilität zwischen Textsignifikaten und Bildsignifikanten muss vorhanden sein. Es gibt also die Möglichkeit, dass ein Bild und ein Text nicht zusammenpassen. Zum einen kann das im Bild visuell Wahrgenommene durch den Text ignoriert werden. Beansprucht der Text jedoch, Signifikate zu bezeichnen, die sich mit solchen des Bildes decken, ohne dass dies zutrifft, wird er durch das Bild falsifiziert. Allerdings wird der Betrachter in diesem Falle versuchen, entweder diese Differenz selbst als signifikant oder die primären Bildsignifikate als Signifikanten einer tropischen Rede zu interpretieren.¹⁶

Ein Bild sagt also nicht nur mehr als tausend Worte,¹⁷ es lässt sich schwerer widerlegen. Ob dies als Vor- oder Nachteil gewertet wird, hängt wohl von den jeweiligen Umständen ab.

14 Ebd. S. 382.

15 Ebd.

16 Ebd.

17 Neben der herkömmlichen hinter dieser Alltagsweisheit stehenden Aussageintention („Ein Referent lässt sich durch ein Bild besser vergegenwärtigen als durch tausend Worte.“), soll hier, angesichts der erfolgten Ausführungen, eine weitere hinzugefügt werden: „Tausend Worte tragen weniger mögliche Bedeutungen als ein Bild.“