

UNIwersYTET IM. ADAMA MICKIEWICZA W POZNANIU
Studia Germanica Posnaniensia XXXV (2014)

KAROLINA KĘSICKA

Uniwersytet im. Adama Mickiewicza,
Poznań

BILDER MIT WORTEN GEMALT – ZUR PROBLEMATIK DER AUDIODESKRIPTION ALS EINER FORM DER ÜBERTRAGUNG VON KUNST

ABSTRACT. Pictures painted in words – about audio description for the Visual Arts

In this paper I shall analyze the audio description for purposes of fine arts as a special form of a multimodal text. With reference to theoretical linguistic concepts of picture as a semiotic sign with iconic but also prototypical function (as a mental picture) this paper shows the differences between visual and verbal signs and attempts to clarify the definition of multimodal competence and their significance for audio describing. On that grounds we examine empirically the selected problems of audiodescription related to composition to demonstrate the difficulties in bringing the visual and verbal components together in one text.

KEYWORDS: audio description for purposes of fine arts, multimodality, mental picture, icon, composition

Die Audiodeskription ermöglicht es blinden und sehbehinderten Menschen, Hörfilme, Theaterstücke und Kunstwerke der bildenden Kunst zu rezipieren. Im Unterschied zu multimedialen Werken wie Film oder Theaterstück, bei denen das Visuelle durch das Auditive (Dialoge, Musik oder Hintergrundgeräusche) unterstützt wird, wird das Visuelle im Falle der Beschreibung von unbewegten Bildern meist nur verbal kompensiert. Nicht selten stellt die Beschreibung durch den Audiodeskriptor den einzigen Weg dar, ein Gemälde oder eine Skulptur für blinde, im Laufe ihres Lebens erblindete und sehschwache Personen rezipierbar zu machen. Von der Präzision der Beschreibung hängt daher ab, wie die Bildwahrnehmung beim Rezipienten entwickelt und gesteuert wird. Für die Ausformulierung der akustischen Gemäldebeschreibung gelten zwar Regeln, die für die Audiodeskription im Allgemeinen gültig sind, ihre Anwendung erweist sich aber wegen der fehlenden Multimedialität als problematischer. Ein Audiodeskriptor versteht sich als ein objektiver Beschreiber, der einem Rezipienten grundsätzlich keine Assoziationen oder Interpre-

tation aufzwingen soll. Bei der Gemäldebeschreibung tritt er allerdings auch in die Rolle des Erlebnis-Vermittlers, der allmählich die Bildkonstruktion vollbringt, indem er die visuellen und kognitiven Vorgänge, die die optische Wahrnehmung begleiten, nachvollzieht und sie dann durch die entsprechende Schematabildung verbalisiert. Durch das Zusammenbringen von Elementen wie Bildkomposition und -technik, Farbe oder Licht entsteht eine komplexe Beschreibung, die aber auch subjektive Züge trägt, da die Auswahl von Beschreibungsdetails eine individuelle und autonome (d.h. durch keine anderen Medien außer dem visuellen Kanal determinierte) Entscheidung des Beschreibers bleibt.

In diesem Zusammenhang setzt sich der vorliegende Beitrag zum Ziel, der Frage nach eigenspezifischen Problemen der akustischen Beschreibung eines Kunstwerks als Form der Audiodeskription nachzugehen. Ins Blickfeld fällt hier u.a. folgende Fragestellung:

1. Lassen sich alle Elemente der Bildkomposition (z.B. Farb- und Lichtkontraste, Duktus, Relativität der Farbwahrnehmung) sprachlich wiedergeben?
2. Wie unterscheiden sich ein bewegtes und unbewegtes Bild voneinander und wie wird die Dynamik eines statischen Bildes bildtechnisch und verbal erzielt?
3. Ist Sprache als Medium genügend präzise, um ohne Hinzuziehung anderer Wahrnehmungskanäle eine entsprechende Schematabildung (das Hervorrufen gleicher mentaler Bilder, die Rekonstruktion sphärischer oder geometrischer Beziehungsverhältnisse und ihr simultanes Zusammenwirken im Akt der Bildperzeption und -rezeption) zu ermöglichen?
4. Ist eine auditive Beschreibung eines Kunstwerks tatsächlich objektiv?
5. Ist die Bildnarration linear rekonstruierbar? Folgt sie überhaupt einem linearen Schema?
6. Wie wird die Bildwahrnehmung durch die zunehmende Detaillierung der Beschreibung gesteuert?

Die Liste von forschungsrelevanten Aspekten ist damit sicher noch längst nicht erschöpft und kann in einer bildlinguistischen, semiotischen oder übersetzungstheoretischen Analyse relevante Anhaltspunkte für eine Vervollständigung und Verifizierung finden. Angesichts der Tatsache, dass diese Form der Audiodeskription als theoretisches sowie empirisches Forschungsfeld noch relativ unterrepräsentiert ist, versteht sich auch dieser Beitrag zwangsläufig eher als theoretische Fragestellung denn als Antwortfindung. Zur besseren Veranschaulichung von Einzelproblemen wird im Anschluss an die theoretischen Überlegungen zu ausgewählten Beispielen der akustischen Gemäldebeschreibungen aus den Sammlungen der Polnischen Nationalgalerie in Warschau übergegangen, um anschließend auf dieser Grundlage Rückschlüsse über die Relevanz der Problemdarstellung ziehen zu können.

Zur Bildtheorie

Bevor auf die Einzelfragen der akustischen Beschreibung von Kunstwerken eingegangen wird, ist zunächst auf einige Begrifflichkeiten hinzuweisen, die für die Audiodeskriptionserforschung einen relevanten theoretischen Ausgangspunkt darstellen. Gemeint sind hier in erster Linie Begriffe, die im Bereich der Bildlinguistik untersucht werden, wie Bildbegriff, multimodale Kompetenz, Bildverstehen oder Bildlichkeit. Sie sind bereits zum Gegenstand theoretischer Reflexion in Bezug auf Hörfilme (vgl. u.a. Fix 2011: 305–329, Seiffert 2005: 67–86, Poethe 2005: 33–48, Stöckl 2004) geworden. Wie bereits erwähnt, sind Forschungsarbeiten zur Audiodeskription eines Kunstwerks aber nicht sehr weit vorangeschritten und es besteht zweifelsohne ein Bedarf, diese Defizite zu decken. In der sich immer stärker digitalisierenden Informationswelt wird es nämlich möglich, Blinden und stark Sehbehinderten immer mehr Inhalte zugänglich zu machen. Damit sich aber dieser Prozess aus pragmatischer Sicht als erfolgreich erweisen kann, müssen wir uns auch theoretisch in die Natur der Dinge vertiefen. Bevor ein Bild sprachlich transkribiert wird, muss man sich also der Spezifik bildlicher Zeichen bewusst werden. Dies können theoretische Ansätze leisten.

Beginnen wir mit dem weitgefassten und inhomogenen Bildbegriff. Bilder werden wissenschaftlich im Sinne materieller Abbildung von Ähnlichkeiten, also in ihrer mimetischen Funktion betrachtet, aber auch als nicht materielle Zeichenkomplexe, die nur gedachte Vorstellungen über ein reales oder fiktives Objekt darstellen. Aus kunstgeschichtlicher und -wissenschaftlicher Sicht sind Bilder eine künstlerische Ausdrucksform und können somit im Zeitalter des pictorial turns als nonverbale Kommunikationsakte gelten. Von der Seite der linguistischen Analyse von Bild-Text-Zusammenhängen können Bilder schließlich als „Komplement eines Sprachtextes“ (Fix 2011: 305), also ein „Element des gesamten Verstehensangebotes des betreffenden Textes“ aufgefasst werden (Fix 2011: 305).

Wie Klemm und Stöckl (vgl. Klemm/Stöckl 2011: 7–9) bemerken, zeigen sich Bilder als Forschungsobjekte in der Hinsicht problematisch, dass ihre Forschung über mehrere Jahre einzelwissenschaftlich betrieben wurde, während heutzutage ein interdisziplinärer Zugriff als angemessener erscheinen muss. Die Autoren verweisen auch auf einige Schwachstellen linguistischer Bildanalyse, die ihres Erachtens behoben werden sollten. Es wird v.a. postuliert, davon abzugehen, anstelle von Bildern als homogenem Begriff über konkrete Bildtypen zu sprechen und sie auf mögliche Formen und Funktionen zu untersuchen (vgl. Klemm/Stöckl 2011: 9). Dies trifft beispielsweise auf die Unterscheidung zwischen bewegten und unbewegten Bildern zu, die keineswegs pauschal als in dieselbe Kategorie fallend eingestuft werden dürfen. Bemängelt wird außerdem, dass Bilder linguistisch oft in Isolation behandelt werden, was sich im Verhältnis zum Forschungsstand anderer Disziplinen als obsolet erweist, seitdem sich disziplinenübergreifend die Einsicht durchgesetzt hat, dass „all media are

mixed media, and all representations are heterogenous; there are no ‚purely‘ visual or verbal arts“ (Mitchell 1995: 5). Diesem Postulat scheint die Audiodeskription als Forschungsdisziplin wohl bereits gewachsen zu sein, indem sie in ihrer Grundannahme von der Multimodalität des Textes ausgeht und das kommunikative Zusammenwirken von Visuellem, Sprachlichem und Auditivem sowohl theoretisch untersucht als auch in der Praxis bei der Erstellung von Hörtexten stets beachtet.

Von mehreren Grundfragen, die sich Klemm, Stöckl zufolge im Hinblick auf Multimedialität von Texten für die Bilduntersuchung als zentral erweisen, sind in Bezug auf die Untersuchung von Hörtexten im Besonderen folgende zu nennen (Klemm/Stöckl 2011: 12):

- „Wie funktionieren Bilder als Zeichen? Wie erkennen wir in ihnen Struktur, Sinn und Zweck?“
- „Wie ist das Verhältnis von Bildinhalt zu formaler Bildgestaltung?“
- „Was unterscheidet Sprache und Bild grundsätzlich? Nach welchen Prinzipien und Mustern geschieht die Verknüpfung von Sprache und Bild? Wie unterscheiden sich dabei statische und bewegte Bilder?“
- „Welche Kognitionsleistungen müssen die Rezipienten von Sprache-Bild-Texten erbringen?“

Antworten auf einige der genannten Fragen sind u.a. in theoretischen Ansätzen (wie Prototypensemantik, Kognitionswissenschaft, kognitive Semiotik oder funktionale Grammatik) zu suchen, die sich Bildern im Sinne ihrer prototypischen und nicht prädikativen (ikonischen) Funktion zuwenden. Veranschaulichen ist offensichtlich eine Primärfunktion von Bildern. Visualisierung kann allerdings nicht der Vermittlung von Bedeutungen gleichgesetzt werden. Durch materielle Bilder wird lediglich die visuelle Charakterisierung eines Objekts erreicht. Die thematische Charakterisierung, also die Zuweisung von Inhalt und Bedeutung, erfolgt erst in mentalen Prozessen, die durch das Betrachten des jeweiligen Bildes in die Wege geleitet werden und bei denen bereits erworbene Wissensbestände und eingeprägte Verknüpfungsmuster wachgerufen werden. Man spricht in diesem Kontext von mentalen Bildern. Diese wirken prototypisch in dem Sinne, dass sie eine typische geistige Vorstellung von einer Sache hervorrufen. Das Evozieren von mentalen Bildern heißt aber nicht gleich das Aufzwingen eines standardisierten Interpretationsschemas. Dass Sprache und Bilder an typische geistige Begriffsvorstellungen anknüpfen, bedeutet noch längst nicht, dass verbal und ikonisch stimulierte kognitive Operationen wie Assoziieren, Vergleichen oder Einordnen immer zu denselben Sinndeutungen führen werden. An mentalen Bildern bleibt (entsprechend der Auffassung: Keine Bedeutung ohne Interpret) immer ein Element der Subjektivität haften. Die Kompetenz des Menschen, bei der Bildwahrnehmung die oben genannten mentalen Prozesse zu initiieren, wird auch multimodale Kompetenz bzw. Multimodalität genannt und als

- „eine gesamtkulturelle Kompetenz und eine individuelle Intelligenz, die das Erschließen von Sinn durch ‚Transkribieren‘, d.h. Kommentieren, Erklären oder Paraphrasieren von Botschaften eines Zeichensystems/Mediums durch ein anderes beinhaltet“ (Stöckl 2011: 47), bzw. als

- „eine Regeln folgende und Muster bildende semiotisch-kognitive Aktivität in der Produktion und im Verstehen von Texten“ (Stöckl 2011: 47) bezeichnet.

Die multimodale Kompetenz ist also eine Fähigkeit, Zusammenhänge zwischen den Modi Sprache und Bild zu erkennen und sie in den Prozess des Gesamtverstehens zu integrieren.

Das kognitionswissenschaftliche Konzept des mentalen Bildes konzentriert sich somit auf die „Bildlichkeit der inneren Vorstellungsinhalte“ (Steinbrenner/Winko 1997: 29) und birgt – wie Fix zu Recht bemerkt – relevante Erkenntnisse für die Entwicklung der Audiodeskription in sich. Mentale Bilder stellen in der Audiodeskriptionsarbeit eine wichtige Bezugsgröße dar, da Audiodeskriptoren nie außer Acht lassen, dass die mentale Bildwahrnehmung bei Sehenden und Nichtsehenden oft unterschiedlich erfolgen kann (vgl. Fix 2011: 310). Aus dieser Reflexion ergibt sich dann für Fix die Fragestellung, inwieweit die Sprache imstande ist, der inneren Vorstellung von Bildern Folge zu leisten. Dies führt für sie zu der Schlussfolgerung, dass ein Identitätsverhältnis zwischen den gegebenen realen oder fiktiven Objekten und ihrer sprachlichen Charakterisierung kaum herzustellen ist. Damit wird an den Begriff der sprachlichen Bildlichkeit angeknüpft, also an das Potenzial der Sprache, Bilder zu generieren. Obwohl die sprachliche Bildlichkeit in der Hinsicht ihre Limits aufweist, lässt sie sich grundsätzlich auf dreierlei Weise evozieren: durch Wortsemantik (Deutungspotenzial der Sprache), Ikonisierung von Sprachzeichen und Metaphorisierung von Inhalten durch die Sprache (vgl. Fix 2011: 309). Auf die Audiodeskription bezogen, ergibt sich daraus die Frage, wie sich das limitierte Bildlichkeitspotential der Sprache auf die Rekonstruktionstreue von Bildern auswirkt. Dass in keinem Übertragungsprozess Inhaltsverluste zu vermeiden sind, ist offensichtlich. Wichtig ist allerdings das Resultat von semantischen und inhaltlichen Verschiebungen. Auf die Audiodeskription eines Kunstwerks trifft dies besonders stark zu, da ein Textrezipient hier keine Möglichkeit hat, die fehlenden Inhalte selbst zu revidieren.

Der Textrezipient ist als Bildbetrachter allein auf eine schriftliche Vertextung angewiesen, deren Wahrnehmungsweise sich wesentlich von der Wahrnehmungsweise des Bildlichen unterscheidet. Während das Schriftliche sukzessiv und linear wahrgenommen wird, perzipiert man Bilder vom Ganzen zu den Teilen hin. Was zählt, ist der erste Gesamteindruck vom Bild. Aufgrund der perzeptiven Bewertung des Ganzen als Gestalt wird dann auf die Deutung von einzelnen Teilen der Bildkomposition geschlossen. Der Gesamteindruck steuert die weitere Bildrezeption, die sich linear als Abfolge von perzeptuellen und kognitiven Tätigkeiten entwickelt. Bei der weiteren Bildbetrachtung wird unser Blick durch Einzelheiten gefesselt, die man in die Komposition einzuordnen versucht. Auf der Grundlage unseres Vorwissens und unserer Annahmen, die im Laufe des Rezeptionsaktes verworfen bzw. bestätigt werden, wird allmählich das Bildverstehen erschlossen. Die Bildwahrnehmung ist

durch Simultanität (Non-Linearität) gekennzeichnet, ihr Verstehen allerdings ist ein dynamischer Prozess, in dem Bildelemente identifiziert, hierarchisiert und selektiert werden, wobei ihre Kontexterkenntnis und der Wissensstand des Bildrezipienten über diese Elemente sich fortlaufend verändern. Bilderrekonstruktion erfolgt also durch das Aktivieren der Bedeutungspotenziale von Bildern, welches als mentale Operation sich zwar parametrisieren lässt, aber trotzdem immer individuell unterschiedlich abläuft und als Prozess keine feste Reihenfolge aufweist. In welcher Reihenfolge Bildelemente wahrgenommen sowie nach welchem Schema sie als mentale Bilder abgerufen und dekodiert werden, liegt an dem jeweiligen Bildbetrachter. Dass Bildrekonstruktion so unterschiedlich ausfallen kann, ist auch darauf zurückzuführen, dass Bilder grundsätzlich durch eine semantische Dichte gekennzeichnet sind, die aus dem Zusammenwirken von Inhalt und Formelementen resultiert. Im Gegensatz zu sprachlichen Zeichen lässt sich ihre Bedeutung kaum fest bestimmen. Sie bleiben immer vage und unterdeterminiert und dadurch für Individualbewertung anfällig.

Bilder verfügen auch über keine Grammatik im Sinne eines Systems von Regeln, die den einzelnen Bildelementen ihre funktionale Rolle zuweisen, sondern über eine flächig angeordnete Struktur. Aus der Anordnung einzelner Bildelemente auf der Sehfläche sowie anderer visueller Komponenten (Farbe, Linien, Licht, Kontraste, Symmetrie, Proportionen usw.), die sich gegenseitig kontextualisieren, ergeben sich dann formale und inhaltliche Zusammenhänge, ähnlich wie aus Linearität und grammatischen Mitteln bei Schrifttexten.

Audiodeskription eines Kunstwerks als multimodaler Text

Der Spezifik der ‚Grammatik‘ von Bildern hat auch die Audiodeskription zu folgen und sie tut es. Entsprechend der visuellen Perzeptionsweise ist den Richtlinien zur Erstellung von audiodeskriptiven Bildbeschreibungen zufolge (vgl. u.a. Dosch/Benecke 1997; Szymańska/Strzymiński 2010; Szymańska 2014: 13–24) von einer allgemeinen Bildbeschreibung (d.h. der lokalen und zeitlichen Verortung von Schauplätzen und Personen sowie ihrer Handlungen) zur Vermittlung weiterer Details, die den Hintergrund einer Szene darstellen und die Bildmessage inhaltlich vervollständigen lassen, überzugehen. Die zunehmende Detaillierung der Beschreibung dient dazu, sich als Bildempfänger ein verständliches und klares Bild von Visuellem zu machen. Die Skizzierung des Allgemeinbildes bildet somit eine Basis zur allmählichen Erschließung von wechselseitigen Schlüsselrelationen zwischen Bild und Text, die u.a. in Komplementarität, Supplementarität oder Integration (mehr dazu Bucher 2011: 129) bestehen können. Aufgrund von vermittelten Details soll der Bildrezipient imstande sein, eigenständig aus der Kombination von Einzelementen des Textes auf die Struktur und Interpretation des Textes als Ganzen schlie-

ßen zu können. Der Aspekt der Ungesteuertheit des Prozesses der Bedeutungszuweisung wird bei der Erstellung von Audiodeskriptionstexten besonders stark hervorgehoben, denn abgesehen von Film- oder Werkautoren sollen keine Instanzen mehr präsent sein, die im Verstehensprozess mitagieren. Ein Audiodeskriptor tritt im Text als objektiver Beschreiber auf, der Bildrezipienten keine Interpretation aufzwingt. Im Falle von Hörfilmen lässt sich die Objektivität der Beschreibung zum Teil nachprüfen, weil der multimodale Filmtext vielerorts durch Auditives komplementiert wird, das das Verständnis von Bild-Text-Relation erweitert und mitsteuern lässt. Dialoge, Musikbegleitung oder Hintergrundgeräusche sind wichtige Elemente der Filmkomposition, dank denen sich u.a. die non-verbale Kommunikation im Filmtext vollziehen kann und die beim Evozieren von mentalen Bildern und konnotativen Inhalten eine wichtige Rolle spielen. Aus der Integration einzelner Modi, ihrer gegenseitigen Kontextualisierung und dem darauf aufgebauten Syntheseprozess von Teilverständnissen ergibt sich dann das Verstehen des Textes als eines Ganzen.

Im Falle der Audiodeskription eines Kunstwerks haben wir es mit der Beschreibung eines statischen Bildes und mit einem Text zu tun, der meist erst im Moment der Erstellung einer Hörbeschreibung multimodal wird. Dem Textautor fällt nun die Aufgabe zu, die Bild-Text-Bezüge erstmals zu konstruieren, anstatt sie zu rekonstruieren, wie es bei der Filmaudiodeskription der Fall ist. Er kann sich angesichts dessen nicht nur als purer Bildbeschreiber verstehen, sondern muss auch in die Rolle eines Erlebnis-Vermittlers schlüpfen. Die Notwendigkeit dieser Doppelfunktion ergibt sich aus der Tatsache, dass für Blinde und stark Sehbehinderte ästhetische Werte eines Kunstwerks bei fehlender Unterstützung durch andere Modi einfach nicht rezipierbar sind. Der Audiodeskriptor als erster Bildempfänger nimmt ein Kunstwerk optisch als Ganzes wahr, selektiert die Bildelemente im Hinblick auf ihre Relevanz für die Bildkomposition und versucht anschließend die sinnliche Wahrnehmung in eine verbale Schematabildung umzuwandeln. Kurzum, sein Bilderlebnis wird kognitiv verarbeitet und anschließend verbalisiert. Da statische Bilder eine nicht-lineare Struktur aufweisen und keine Stützelemente (wie beispielsweise die Filmhandlung) besitzen, die die Reihenfolge der Beschreibung determinieren würden, bleibt es dem Audiodeskriptor alleine überlassen, zu entscheiden, welche Elemente der Bildkomposition berücksichtigt und in welcher Reihenfolge sie präsentiert werden. In diesem Moment kommt der Aspekt der Autonomie und Subjektivität bei der Auswahl von Bildelementen zum Tragen. Unter Autonomie ist hier zu verstehen, dass die Auswahl weder durch den Zeitfaktor wie bei Spielfilmen oder Theateraufführungen noch durch intermediale Textbezüge determiniert wird. Mit der Autonomie der Auswahl geht zwangsläufig ihre Subjektivität einher, die allerdings aufgrund der Monomodalität von Kunstwerken m.E. die Bildinterpretation viel stärker als bei anderen Formen der Audiodeskription lenken kann.

Während die semantisch-inhaltliche Kohärenz von Audiodeskriptionstexten grundsätzlich durch das Ablesen von multimodalen Textzusammenhängen und im

Falle einer auditiven Kunstwerkbeschreibung durch ihre Konstitution hergestellt wird, ist ihre formale Kohäsion an der strukturellen Anordnung der Beschreibung abzulesen.

Auditiv Kunstwerkbeschreibungen verfolgen als Text folgendes Schema:

- einleitende Informationen über das Kunstwerk (Form, Titel, Name und Vorname des Autors, Entstehungszeit bzw. – Raum, Technik, Maße, Standort)
- die Allgemeinbeschreibung des Kunstwerks (Bildthema, Bild- und Farbkomposition)
- die detaillierte Beschreibung des Kunstwerks (mit einer abschließenden Information über die Signatur des Künstlers)
- Zusatzinformationen über Stil und verwendete Techniken (möglicherweise auch historischer und kunstgeschichtlicher Hintergrund, Anekdoten über Werk oder Autor) (vgl. Szymańska 2014: 13–24).

Wie bereits erwähnt, stellt die Allgemeinbeschreibung den Ausgangspunkt für die detaillierte Beschreibung der Bildkomposition dar. Anders als bei anderen Formen der Audiodeskription, bei denen die Beschreibungsnarration durch die lineare Abfolge von Textsequenzen gesteuert wird, muss der Bildbeschreiber bereits in dieser Phase den Narrationsablauf festlegen, der weiterhin konsequent zu verfolgen ist. Somit wird auch die Bildinterpretation bei Rezipienten konsequent durch die vom Beschreiber subjektiv vorgenommene Informationsselektion gelenkt. An dieser Stelle knüpfen wir wieder an die eingangs gestellte Frage nach der Objektivität der Beschreibung an und stellen fest, dass sie bei der Audiodeskription von Kunstwerken zwar immer intendiert wird, aber nie ganz erreicht werden kann, da die Beschreibung von Kunstwerken neben den Informationen über die logische Anordnung der Bildelemente und der Vermittlung von Fakten sowie rein perzeptiven Eindrücken die verbale Wiedergabe der Bildidee und die Interpretation von Bildkomposition erforderlich macht.

Mit der Erhöhung des Detailliertheitsgrads der Bildwahrnehmung wird die Bildbeschreibung komplexer und inhaltlich dichter. In dieser Phase soll der Bildrezipient imstande sein, aus den gewonnen Einzeldaten die erzählerische Aussage eines Kunstwerkes als Vorstufe seiner eigenen Interpretationsbildung zu rekonstruieren und nachzuvollziehen. Es zeigt sich aber, dass es viel problematischer ist, in Bezug auf die mimetische Wiedergabe von Bilddetails die Inhaltstreue zum Original zu erzielen, als man anfangs vermutet. Probleme ergeben sich grundsätzlich aus zwei Faktoren, und zwar aus der Bildkomposition selbst und aus dem Potenzial des Sprachlichen, Visuelles in mentale Bilder umzusetzen.

Inhalt-Form-Bezüge in der auditiven Beschreibung von Kunstwerken

Der Umfangsrahmen der vorliegenden Arbeit macht es unmöglich, auf alle potenziellen Probleme genauer einzugehen, die mit der Erstellung von Audiodeskriptionstexten in Museen zusammenhängen. Der Schwerpunkt liegt somit darauf, auf

bildkompositorische Zentralfragen hinzuweisen und sie mit einigen Beispielen von Audiodeskriptionen zu kommentieren.

Probleme betreffen v.a. die **Rekonstruktion der Inhalt-Form-Bezüge** sowie der Relationen und Proportionen der Figuren und Gegenstände auf der Fläche (Figur-Grund-Beziehungen). Die Komposition eines Kunstwerkes umfasst die Organisation der Bildfläche, insbesondere die Anordnung der Motive. Sie ist allerdings mehr als ein Resultat des Gestaltungsprozesses, sondern zugleich eine künstlerische Aktivität der Ideenvermittlung mithilfe einer elementaren Formengrammatik. Der Aufbau eines Kunstwerkes ist formal als Summe aller Komponenten und semantisch als bildhafte Idee des Künstlers aufzufassen. Damit Einzelelemente der Bildkonstruktion den künstlerischen Intentionen gemäß als Ganzes nachvollziehbar werden, ist die Kenntnis der Formengrammatik erforderlich. Damit sind künstlerische Gestaltungsmittel, Materialien und Techniken gemeint, insbesondere geometrische Ordnungsprinzipien sowie Elemente wie Rhythmus, Farbgestaltung und Beleuchtung. Sowohl für den Audiodeskriptor als auch den antizipierten Bildrezipienten erweist sich die formale Bildanalyse als problematisch, wenn man von der natürlichen Annahme ausgeht, dass ihre kunsttheoretischen und -geschichtlichen Fachkenntnisse geringer als die des Künstlers sind. Künstler haben theoretisches Wissen über die Bildgeometrie, Bildbetrachter nicht unbedingt. Daher ziehen alle Audiodeskriptoren bei der Texterstellung die Fachberatung von Kunsthistorikern zu Rate. Die Bildkomposition wird für viele Betrachter oft nicht direkt, sondern erst aufgrund von kunstgeschichtlichen Analysen ersichtlich. Fehlt neben dem Fachwissen auch das Sehvermögen, dann muss man sich als Rezipient völlig auf das eigene und fremde Vorstellungsvermögen stützen. Daher ist es bei der Schilderung der Komposition eines Bildes unabdingbar, dass die Beschreibung eine konsequente Strukturiertheit hat, die das Vorstellungsvermögen entsprechend unterstützt.

Besonders schwierig gestaltet sich die Verbalisierung der Inhalt-Form-Struktur bei Bildern, die eine **offene Bildkomposition** aufweisen, d.h. bei solchen, die keine klassischen Kompositionsregeln befolgen. Bei einer individuellen Komposition wird im Vergleich zur mimetisch orientierten klassischen Komposition die **künstlerische Subjektivität** noch verstärkt. Das Bildthema wird hier nur im Ausschnitt präsentiert. Seine Interpretation geht dadurch weit über den Bildraum hinaus. Aus dem im Detail betrachteten Motiv ist dann auf die Komposition als Idee eigenrezeptiv zu schließen. Individuelle Kompositionsformen (wie z.B. der Kubismus) verwenden auch keine gängigen Farbkontraste sowie keine klassische Bildaufteilung und Formenanordnung. Ein Resultat ist dann beispielsweise, dass Gegenstände absichtlich ungeordnet, sogar chaotisch über die Bildfläche gestreut sind. Ein kompositorisches Chaos strukturiert in der auditiven Beschreibung wiederzugeben, muss eine große Herausforderung darstellen.

Als Beispiel für einen die klassischen Prinzipien sprengenden Kompositionsrahmen kann die Bildkomposition im Bild *Akt z kotem* (1920) von Tytus Czyżewski

gelten. Der Körper wird in diesem Frauenakt zerstückelt, wobei die einzelnen Fragmente kompositorisch kein einheitliches Ganzes bilden. Hinzu kommt noch, dass Körperteile der Figur durch andere geometrische Formen in der Fläche verdeckt werden, wodurch die räumlichen Bezüge zwischen einzelnen Bildelementen noch unsichtbarer werden. Auf der anderen Seite sind gerade diese horizontal und vertikal gelegenen Formen Elemente, die die Bildkomposition strukturieren. Durch die **Aufhebung des einheitlichen Blickpunkts** und die **räumliche Fragmentarisierung** wird es allerdings schwierig, topographische Platzierung präzise in der Audiodeskription wiederzugeben.

Zur Bildkomposition überleitend, verweist der Audiodeskriptor auf die Zerstückelung der zentral gelegenen Figur und fügt hinzu, dass die Figur in ein Schema von vertikalen und horizontalen Linien eingebaut wird. Es fehlt jedoch eine Information über die Struktur dieses Schemas, wodurch die Informationen unklar bleiben. Erst in der detaillierten Beschreibung wird auf diese Struktur näher eingegangen, indem zu einem Vergleich der geometrischen Formen mit Vorhängen gegriffen wird. Der Audiodeskriptor bedient sich dieses Vergleichs konsequent und versucht durch die Assoziation mit dem Bühnenbild die Funktion und Lage der räumlichen Elemente sprachlich zu visualisieren. Er knüpft hier an ein vorgeprägtes Begriffsverständnis an, um damit einem Bildrezipienten die Rekonstruktion der formalen Anordnung zu erleichtern. Der Bildraum wird zu einer Bühne, auf der einzelne Bildelemente entweder an die Kulissen oder in den Vordergrund gerückt werden. Trotzdem ist die entsprechende Schematabildung nur partiell vollzogen. Der Grund dafür liegt an der Verbalisierung dieses Schemas, genauer gesagt an der fehlenden Präzision der Sprache. Sprache versteht sich bei auditiver Beschreibung als ein Werkzeug, das einem blinden oder sehbehinderten Bildrezipienten bei der Bildrekonstruktion helfen kann. Daher scheint es geboten, bildhafte und zugleich präzise Lexeme und einfache Satzkonstruktionen zu verwenden, um damit das Vorstellungsvermögen des Rezipienten entsprechend zu aktivieren. Die Frage ist, ob Sprachliches in Audiodeskriptionen in der Tat so bildhaft und präzise sowie interpretationsfrei wirkt, wie gemeinhin erwartet wird.

Bei der Schilderung des Allgemeinbildes wird die Bildrezeption von Anfang an falsch gelenkt, indem durch die Mehrdeutigkeit der Formulierung „obraz przedstawia rozczłonkowane ciało kobiety i kota“ (dt. ‚Das Gemälde stellt einen zergliederten Körper einer Frau und einer Katze dar‘) auch die Segmentierung des Katzenobjekts suggeriert werden kann. Das Sprachliche versagt auch bei der Schilderung der Proportionen. Durch die Information „sylwetka zajmuje całą szerokość obrazu“ (dt. ‚die Figurengestalt ist bildflächenweit situiert‘) erfährt man zwar etwas über die zentrale Lage der Figurengestalt, aber nichts über proportionale Figur-Grund-Verhältnisse. Fraglich bleibt auch die sprachliche Darstellungsweise der Katzenfigur als eines sich bewegenden Objekts („[...] głaszcząc kota, który się przechadza“ dt. ‚[...] streichelt die gehende Katze‘), während sie in Wirklichkeit eher statisch wirkt.

Eine scheinbar **chaotisch gestreute Bildkomposition** durch eine strukturierte lineare Narration in der Audiodeskription wiederzugeben, wird umso schwieriger, je dichter Bildelemente miteinander verflochten sind, wie es im Bild *Miasto fabryczne* (1920) von Leon Chwistek der Fall ist. Aus dem Zusammenspiel von vertikal, horizontal, diagonal und kreisförmig verlaufenden Linien und Konturen ergibt sich die Dynamik der Darstellung, die ein Audiodeskriptor wiederzugeben hat. Hinzu kommt noch, dass sich in der Bildfläche keine deutliche Trennlinie zwischen Mitte und Peripherie ziehen lässt. Ähnlich wie bei Czyżewski setzt sich das Sujet auch hier imaginär über den Bildraum hinaus fort, wodurch einem Bildrezipienten ein erzählerischer Überblick entzogen wird. Als problematisch erweist sich neben der **Kombiniertheit von teilweise aufgelösten, aber überwiegend betonten Umrisslinien** auch das **Bildthema** selbst: die Vorstellung einer Stadt der Zukunft. Die Komposition stellt dabei kein Abbild einer bestehenden Stadtmetropole, sondern die Visualisierung von Moderne dar. Da Moderne als mentales Bild wandelbar ist, kann sie je nach Epoche und individueller Weltwahrnehmung unterschiedlich konnotiert werden. Um dieser begrifflichen Variabilität gerecht zu werden, wird in der Audiodeskription an ein anderes mentales Bild angeknüpft, und zwar an das pop-kulturelle Bild von Gotham City als Stadt der Zukunft, um damit eine Art mentaler Brücke zwischen zeitlich variablen Begriffsvorstellungen zu bauen. Dies ist eine interessante Visualisierungsstrategie, die es ermöglicht, die anfangs vage Begriffsvorstellung zu konkretisieren. Das mentale Bild von einer Stadt der Zukunft wird auch mit dem Bild Gdingens als dem Symbol der Moderne verglichen. Da aber Gdingen seinen Symbolstatus heute völlig verloren hat, scheint diese Schematabildung wenig Chancen zu haben, entsprechende Resonanz zu finden.

Die Verflechtung von Linien, insbesondere die Dominanz vertikaler Richtungen, ergibt den **Rhythmus**, erschwert aber gleichzeitig die Entschlüsselung der Komposition. Im Resultat ist man als Betrachter nur imstande, seine Erlebnisse zu verbalisieren. Ein Audiodeskriptor tritt in die **Rolle des Erlebnis-Vermittlers**, indem er die Komposition seiner subjektiven Perzeptions- und Rezeptionsweise gemäß beschreibt. Sprachlich wird es u.a. durch folgende Formulierungen zum Ausdruck gebracht: „dają złudzenie ruchu“ (dt. ‚täuscht eine Bewegung vor‘), „przypomina krajobraz Gotham City“ (‚ähnelt der Landschaft von Gotham City‘), „wzmaga efekt szkicowości“ (dt. ‚betont den Effekt der Skizzenhaftigkeit‘) oder „zabudowania jakby z poprzedniej epoki“ (dt. ‚Bauten wie aus einer früheren Epoche‘). Alle genannten Formulierungen haben neben der **Betonung der Subjektivität der Wahrnehmung** eine semantische Vagheit gemein. Beides stellt wiederum einen groben Verstoß gegen die Hauptregeln bei der Erstellung von Audiodeskriptionen dar, und zwar Dinge präzise und so zu beschreiben, wie sie sind, ohne sie zu interpretieren. Bei einer solchen **Kompositionsdichte** ist aber ihre treue Wiedergabe in der Audiobeschreibung kaum möglich. Um die Beschreibungsklarheit nicht aus den Augen zu verlieren, muss die Beschreibung fragmentarisch erfolgen. So geschieht es auch im

besprochenen Beispiel, in dem die Beschreibung sich auf die Diagonalen konzentriert und andere Elemente als peripher betrachtet. Durch die Konzentration der Beschreibung um die diagonale Achse gewinnt die Beschreibung an Kohärenz und Übersichtlichkeit, wird aber durch die Auslassung anderer Elemente unvollständig.

Ein weiteres kompositionsbezogenes Problem betrifft **Elemente, die das rein Perzeptive ansprechen**. Licht- und Schattenmodulation, der Zusammenhang von Farben und Licht, Farbwirkung und Relativität der Farbwahrnehmung oder Duktus sind Kompositionselemente, die sich im Prinzip kaum verbal wiedergeben lassen. Selbst eine sehr detaillierte Farbenbeschreibung ist nicht imstande, die **Dynamik der Sinneswahrnehmung zu ersetzen**. Farben, Licht und Malauftrag sind in ihrem gleichzeitigen Zusammenwirken für die Bildführung zuständig. Samt den übrigen Elementen der künstlerischen Organisation der Bildfläche wirken sie ganzheitlich auf den Bildrezipienten und lenken seinen Wahrnehmungsprozess. Farben sind Ausdruck individueller und subjektiver Empfindungen des Künstlers und so werden sie auch von Bildbetrachtern wahrgenommen. Die Versprachlichung von Sinnesempfindungen subjektiviert sie dadurch noch tiefer, dass sie durch Sinneseindrücke eines anderen Betrachters gefiltert werden. Aufgrund der Linearität der Beschreibung geht außerdem die für die Bildbetrachtung so wichtige **Simultaneität der Wahrnehmung** verloren.

Die ziemlich monotone Farbgestaltung im Bild *Akt z kotem* von Czyżewski wird durch einzelne Farbkontraste (v.a. Rot) dynamisiert. Farben, Schatten und Licht werden gezielt schematisch aufgetragen, um dem Bild den Anschein einer technischen Unbeholfenheit des Autors zu verleihen. Diesen Eindruck sollen noch die Malweise und zahlreiche kleine Pinselstriche verstärken. In der Audiodeskription beschränkt sich die Farbenbeschreibung lediglich auf das Nennen von dominanten Farben ohne Präzisierung von Farbtönen. Nicht angesprochen werden auch die **Farbkontraste** sowie der **Duktus**, was zur Folge hat, dass sich ein blinder oder stark sehbehinderter Bildrezipient ihrer Funktion nicht bewusst wird. In Bezug auf die Schilderung des Frauengesichts wird zwar die Formulierung „namalowana schematycznie“ (dt. ‚schematisch gemalt‘) verwendet, aus der ohne Hintergrundinformationen über das konsequente Imitieren einer technisch schwachen Malweise aber nicht auf die Intention des Autors zu schließen ist.

Als schwierig bei der Audiodeskription eines Kunstwerkes erweist sich auch eine Raumgestaltung, die sich auf Sinnesempfindungen stützt und die **empirische Wahrnehmung nachzuahmen** versucht, wie z.B. Notturmi, in denen die Komposition das Dunkelheitserlebnis zu vermitteln hat. Charakteristisch für eine solche Komposition sind also wenig betonte Umrisse, eine stark reduzierte Farbpalette oder sogar ein monochromes Farbkonzept und ein **Zusammenspiel der Monochromatik der Farben und der Lichtkontraste**. Als Beispiel dafür kann das Bild *Opera paryska w nocy* (1898) von Aleksander Gierymski gelten. Das Bild imitiert visuelle Sinneseindrücke, die man bei der Objektbetrachtung im Dunkeln erlebt. Anfangs

wirkt die Bildfläche wie ein Farbfleck ohne zeichnerisches Gerüst. Verschwommene Konturen und die monochrome Farbgestaltung erschweren die räumliche Orientierung. Je länger man jedoch das Bild betrachtet, desto sichtbarer wird die Komposition. So wie sich die Augen an die Dunkelheit gewöhnen, so wird auch der Blick des Rezipienten auf immer mehr Details gelenkt, als würde sich der Bildraum allmählich erhellen. Einen solchen perzeptiven Vorgang beim Rezipienten in Gang zu setzen, ist dank der Modulation von Schwarztönen und Lichteffekten möglich, die für eine entsprechende Spannung und Stimmung im Bild sorgen. Auch wenn sich ein Blinder oder stark Sehbehinderter der Farbwirkung bewusst wird, ist bei ihm der **Effekt der sinngestützten Relativität der Farb- und Lichtwahrnehmung nicht abrufbar**. Helldunkelbeziehungen, die sich durch Licht- und Farbkontraste und Farbumgebung aufbauen, lassen sich sprachlich zwar beschreiben, aber nicht ganz imitieren.

Trotzdem ist anzumerken, dass Sprachliches dank der Metaphorisierung nicht ganz hilflos vor der Aufgabe steht, Sinneswahrnehmungen (hier Farbwahrnehmung) zu vermitteln. Im Audiodeskriptionstext zum Bild *Opera paryska w nocy* wird auf zahlreiche Metaphern zurückgegriffen, die die Farb- und Lichtwirkung beschreiben lassen: „falują blaski odbijającego się światła“ (dt. ‚Lichtreflexe bewegen sich‘), „mieniący się odcieniami błękitu“ (dt. ‚blautönig schimmernd‘), „sylwetka rozplywa się w mroku“ (dt. ‚Die Silhouette verschwimmt im Dunkeln‘), „migające sylwetki“ (dt. ‚schimmernde Silhouetten‘), „tańczą cienie“ (dt. ‚Schatten tanzen‘), „fasada majaczy“ (dt. ‚die Fassade erblinkt‘) u.ä. Der Audiodeskriptor verwendet darüber hinaus die Strategie der kognitiven Rekonstruktion des Wahrnehmungsvorgangs, indem er bei der Allgemeinbeschreibung den Blick zuerst auf die Beschreibung einer Straßenlaterne als eines zentralen Kompositionselements lenkt und von diesem Punkt an zu den übrigen Bildelementen überleitet. Subjektive Sinneseindrücke und Lichteffekte, die dabei entstehen, werden dann von ihm kognitiv rezipiert und kommentiert (z.B. „Nasz wzrok zaczyna się przyzwyczajać do panującej ciemności i zaczyna wyłuskiwać kolejne elementy“ – dt. ‚Unser Auge gewöhnt sich allmählich an die Dunkelheit und beginnt, einzelne Elemente zu erkennen‘, „Z ciemności powoli wyłania się ...“ – dt. ‚Aus der Dunkelheit tritt allmählich ... hervor‘, „trzeba wyteżyc wzrok“ – dt. ‚Man muss seine Augen anstrengen‘), wodurch der Zielrezipient mit der künstlerischen Idee vertraut gemacht wird.

Fazit

Aus den besprochenen Beispielen ergibt sich folgende Schlussbetrachtung: Eine inhaltstreue und ideengerechte Audiodeskription eines Kunstwerkes zu erstellen, stellt eine große kognitive Herausforderung für Bildbeschreiber dar. Der Bildbeschreiber steht vor der Notwendigkeit, ein künstlerisches Konzept sprachlich zu visualisieren. Er bewegt sich also zwischen zwei unterschiedlichen Medien: dem

Visuellen und dem Sprachlichen. Dies erfordert von ihm die Entwicklung einer Sonderkompetenz und zwar der multimodalen Kompetenz. Multimodale Texte, darunter Audiodeskriptionen, sind komplexe Sprach-Bild-Gebilde, bei deren Erstellung mehrere Aspekte zu beachten sind. In Bezug auf audiodeskriptive Beschreibungen von Kunstwerken sind es v.a. räumliche und zeitliche Aspekte der Bildperzeption, die Architektur der Bilder, Rezeptionssteuerung, die Selektion von zu beschreibenden Bildelementen und die Strategieentwicklung bei der Wiedergabe von unübertragbaren Kompositionselementen wie Lichtverhältnissen, Farb- oder Linienwirkung. Als besonders problematisch erweisen sich kompositionsbezogene Elemente, da erstens ihre Beschreibung von Audiodeskriptoren ein kunsttheoretisches Fachwissen erfordert und es zweitens der Sprache oft an Bildhaftigkeit und den Sprachbenutzern an Präzision fehlt, um sie adäquat zu beschreiben. Hinzu kommt, dass in Richtlinien zur Erstellung von Audiodeskriptionen kompositionsbezogenen Aspekten wenig Aufmerksamkeit geschenkt wird. Auch die theoretische Reflexion dieser Problematik, die aus der Perspektive der Praxis sehr aufschlussreich wäre, ist nach wie vor als unzureichend anzusehen.

Dieser Sachverhalt hat zur Folge, dass Audiodeskriptoren von Kunstwerken im Vergleich zu Autoren von Hörfilmen intuitiv vorgehen müssen. Ihre Arbeit trägt auch dadurch subjektive Züge, dass sie autonom über Inhaltsselektion entscheiden und dass sie den Rekonstruktionsprozess in großem Maße auf eigene Wahrnehmung stützen. Dies ist darauf zurückzuführen, dass das Visuelle an mentale Begriffsvorstellungen, Schematabildung und Assoziationen anknüpft und als solches keinen Anspruch auf Objektivität erhebt. Bei der Erstellung einer Audiodeskription werden Bilder perzipiert und anschließend rezipiert. Beides ist subjektiv. Wichtig ist allerdings, dass man sein eigenes Bildverstehen mental verarbeitet, für sich selbst visualisiert und versucht das Bild in einer Weise zu rekonstruieren, die rezipientengerecht ist. Man kann aber nicht garantieren, dass bei blinden oder sehbehinderten Bildrezipienten dieselben mentalen Bilder wie bei sehenden Bildbetrachtern hervorgerufen werden.

Vom Gesichtspunkt nichtsehender oder sehschwacher Bildrezipienten ist die Perzeption und Rezeption multimodaler Texte auch keine leichte Aufgabe. Sie müssen über einen längeren Zeitraum ihre Konzentration entsprechend der Beschreibungsrichtung lenken, auf jeder Beschreibungsetappe die Narration als Ganzes erfassen und beim Hervorrufen von nachfolgenden mentalen Bildern stets kognitive Prozesse der Informationssynthese in Gang setzen.

Abschließend lässt sich pessimistisch feststellen, dass Audiodeskriptionen von Kunstwerken mit einem Mangel behaftet sind, der im Prinzip nicht zu beseitigen ist, da die Beschreibung linear erfolgt, während sich die Komposition als Ausdruck künstlerischer Idee nur bei simultaner Betrachtung aller Elemente als Ganzes offenbart. Dieser Herausforderung wird Sprachliches natürlich nie gewachsen sein. Trotzdem oder gerade deswegen brauchen wir eine vertiefte und mehrdimensionale

Forschung über diese multimodale Textsorte, um die Arbeit der Audiodeskriptoren zu erleichtern und die Qualität entstehender Audiodeskriptionen verbessern zu können.

Literatur

- Audiodeskriptionen zu Bildern Akt z kotem von Tytus Czyżewski, Opera paryska w nocy von Aleksander Gierymski und Miasto fabryczne von Leon Chwistek, <<http://mnw.art.pl/multimedia/audiodeskrpcje/dziela-ze-zbiorow-muzeum-malarstwo>>, (Zugriff am 30.10.2014).
- Bucher, Hans-Jürgen: *Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität*, in: Hajo Diekmannshanke u.a. (Hg.): *Bildlinguistik. Theorien-Methoden-Fallbeispiele*, Berlin 2011, S. 121–156.
- Dosch, Elmar/Bernd, Benecke: *Wenn aus Bildern Worte werden. Ein Handbuch für Filmbeschreiber*, München 1997.
- Fix, Ulla: *Bilder wahrnehmen, ohne zu sehen? Bildlichkeit in der Audiodeskription von Hörfilmen*, in: Hajo Diekmannshanke u.a. (Hg.): *Bildlinguistik. Theorien-Methoden-Fallbeispiele*, Berlin 2011, S. 305–329.
- Klemm, Michael/Hartmut, Stöckl: „Bildlinguistik“ – Standortbestimmung, Überblick, Forschungsdesiderate, in: Hajo Diekmannshanke u.a. (Hg.): *Bildlinguistik. Theorien-Methoden-Fallbeispiele*, Berlin 2011, S. 7–18.
- Mitchell, William J. Thomas: *Picture Theory. Essays on Verbal and Visual Representation*, Chicago 1995.
- Poethe, Hannelore: *Audiodeskription – Entstehung und Wesen einer Textsorte*, in: Ulla Fix (Hg.): *Hörfilm. Bildkompensation durch Sprache*, Berlin 2005, S. 33–48.
- Seiffert, Anja: *Räumliches hören. Ein schemaorientierte Analyse der audiodeskriptiven Darstellung der Handlungsräume*, in: Ulla Fix (Hg.): *Hörfilm. Bildkompensation durch Sprache*, Berlin 2005, S. 67–86.
- Steinbrenner, Jakob/Ulrich, Winko: *Bilder in der Philosophie & in anderen Künsten & Wissenschaften*, Paderborn 1997.
- Stöckl, Hartmut: *Die Sprache im Bild – Das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text*, Berlin 2004.
- Stöckl, Hartmut: *Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz*, in: Hajo Diekmannshanke u.a. (Hg.): *Bildlinguistik. Theorien-Methoden-Fallbeispiele*, Berlin 2011, S. 45–70.
- Szymańska, Barbara: *Rekomendacje dotyczące udostępniania instytucji muzealnych osobom z niepełnosprawnością wzroku i tworzenia audiodeskrypcji do dzieł plastycznych*, <http://nimos.pl/upload/muzea_bez_barier/Rekomendacje_dotyczace_audiodeskrypcji_w_muzeach.pdf>, (Zugriff am 30.10.2014).
- Szymańska, Barbara/Strzyński, Tomasz: *Audiodeskrypcja. Obraz słowem malowany. Standardy tworzenia audiodeskrypcji do produkcji audiowizualnych*, <http://avt.ils.uw.edu.pl/files/2010/12/AD/_standardy_tworzenia.pdf>, (Zugriff am 30.10.2014).